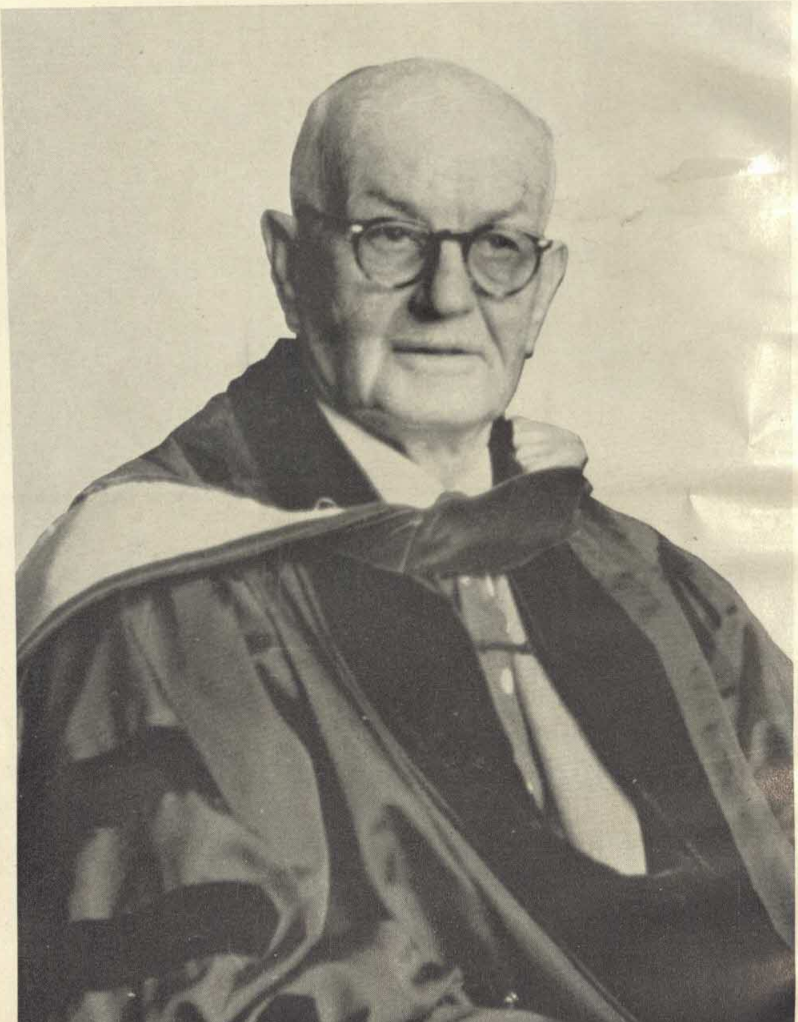


Edward C. Tolman

(See page 135)

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXXIV MARCH, 1961 No. 1

AN INVESTIGATION OF FIGURAL ADAPTATION: A STUDY WITHIN THE FRAMEWORK OF SENSORY-TONIC FIELD-THEORY

By SONOKO OHWAKI, Clark University

Figural adaptation is discussed here in terms of systematic changes in perception that occur when distorting media, such as prisms, are introduced. Such systematic changes in perception are known to be in the direction of reëstablishing an orderly perceptual world. The present study is directed toward an experimental analysis of factors involved in figural adaptation conceptualized from the point of view of an organismic theory, viz. sensory-tonic field-theory. The means of disturbing normal conditions of stimulation used in the present investigation, viz. prisms, have been employed by other investigators. Following Stratton's pioneering work, several investigators have utilized distorting prisms to study adaptation to a distorted field.[1] Quite recently I. Kohler in particular has made some significant advances in the investigation of figural adaptation by introducing prisms

* Received for publication August 16, 1959. This investigation was in part supported by the Research Grant M-348 from the National Institute of Mental Health, National Institute of Health, Public Health Service.

[1] G. M. Stratton, Vision without inversion of the retinal image, *Psychol. Rev.*, 4, 1897, 341-360, 463-481; Margaret Wooster, Certain factors in the development of a new spatial coordination, *Psychol. Monogr.*, 32, 1923, (No. 146), 1-96; G. G. Brown, Perception of depth with distorted vision, *Brit. J. Psychol.*, 19, 1928, 117-146; P. H. Ewert, A study of the effect of inverted retinal stimulation upon spatially coordinated behavior, *Genet. Psychol. Monogr.*, 7, 1930, 177-363; J. J. Gibson, Perception of curved lines, *J. exp. Psychol.*, 16, 1933, 1-31; F. W. Snyder and N. H. Pronko, *Vision with Spatial Inversion*, 1952, 1-144; I. Kohler, Ueber Aufbau und Wandlungen der Wahrnehmungswelt, *Oesterr. Akad. Wissensch., Philos.-Histor. Kl., Sitz-Ber.*, 227, 1951, 1-118; Joseph Bossom, and Richard Held, Shifts in egocentric localization following prolonged displacement of the retinal image, *Amer. Psychologist*, 12, 1957, 454 (abstract); Duilio Giannitrapani, Adaptation to a distorted visual field, *ibid.*, unpublished paper read at E.P.A. meeting, 1957 (abstract); Changes in adaptation to prolonged perceptual distortion: A developmental study, Unpublished doctoral dissertation, Clark University, 1957.

which distort only a portion of the visual field, and glasses which color parts of the visual field differently.

The essential tenet of our approach to figural adaptation is that the perceptual process must be analyzed in terms of total organismic activity rather than in terms of peripheral sensory functioning. Perception is considered here as an experience which corresponds to a particular relation between proximal stimuli (s) and the ongoing state of the organism (o). A further assumption concerns basic s/o relationships in terms of stability-instability. A stable s/o relationship is defined as one where, given a certain proximal stimulus (s), there is *no* tendency for the pertinent aspect of the organismic state (o) to change. In contrast, an unstable relationship is defined as one where there *is* a tendency for the pertinent aspect of the organismic state to change. It should further be noted that the establishment of an orderly perceptual world entails the formation of certain stable s/o relations which correspond to perceptions of particular significance, such as perception of verticality and horizontality in regard to location, and perception of straightness in regard to form. Unstable s/o relationships correspond to perceptions of non-verticality, non-horizontality (tilt), and non-straightness (curvature). Taking into account these notions, adaptation is assumed to involve the following; certain proximal stimuli s_x issuing from a physical stimulus (S)—in regard to location, a plumb line—are under normal conditions in stable relation to organismic state o_x; the proximal stimuli issuing from the same S (plumb line) through a distorting prism have changed from s_x to s_y and are in unstable relation to the organismic state o_x. For figural adaptation to occur, that is, in order for the proximal stimuli s_y to assume the particular perceptual significance of s_x, *i.e.* verticality, the organismic state must change from o_x to o_y.[2]

Consonant with organismic theory, the further assumption is made that a change in organismic state pertains not to one modality alone, but rather to other modalities relevant to the perception with which *S* is concerned. If, for instance, one deals with verticality, one would expect adaptation within the visual modality to have its effect in the tactual kinesthetic perception (intermodal figural adaptation). Analogously, if figural adaptation occurs in one eye, we may expect it to occur in the other eye (interocular figural adaptation).

[2] Heinz Werner and Seymour Wapner, Toward a general theory of perception, *Psychol. Rev.*, 59, 1952, 324-338; Wapner and Werner, *Perceptual Development*, 1957, 1-13.

A second aspect of this study concerns the process of readaptation which we consider fundamentally to be the same as that of adaptation. This part of the investigation is exploratory though certain expectation may be stated. Without going into a theoretical rationale, one should expect that the number and variety of cues available to the organism for exploitation in the process of readaptation should be one of the factors influencing the rate of readaptation. In other words, with more cues available there should be a faster rate of readaptation. One might expect, therefore, that a subject (*S*) exposed to visual stimuli in addition to tactual-kinesthetic stimuli during a period of readaptation, will adapt faster than an *S* exposed to only a limited amount of tactual-kinesthetic stimuli.

General Procedure

Our expectations regarding interocular and intermodal figural adaptations and the different rate of readaptation during reduced stimulation, were tested in two experiments: Experiment I, Adaptation to a tilted visual field; and Experiment II, Adaptation to a curved field. In Experiment I, *S* wore a prism which rotated the total field and figural adaptation was measured in terms of the location of a line which appeared vertical. In Experiment II, *S* wore a prism which convexed all vertical lines to the right and figural adaptation was measured in terms of the curvature of a line which appeared straight. The general procedure for both experiments consisted of seven steps as schematized in Table I. This sequence was as follows. (1) Pretest: A measurement of visual and kinesthetic apparent verticality (Experiment I) or apparent straightness (Experiment II) prior to the introduction of distorting prisms. This is the control condition. (2) Adaptation-period: The *Ss* wore prisms and moved around or sat in a lighted room. (3) Post-test I: A visual and kinesthetic measurement followed the removal of the prisms. (4) First readaptation period: During this period of 3 min. and the following readaptation-period (Step 6 in Table I), one half of *Ss* (Exposed Group) moved or sat in a lighted room while the other half (Unexposed Group) were seated in a dark room. (5) Post-test II: A visual and kinesthetic measurement following the first readaptation-period. (6) Second readaptation-period: A period of 5 min. comparable to Step (4). (7) Post-test III: This step followed the second readaptation-period—a total of 8 min. readaptation. The procedure specific to each step is described in the following section. The *Ss* were 20 undergraduate students (10 men and 10 women).

Experiment I: Adaptation to a Tilted Visual Field

(1) Apparatus. (a) Distorting prism. The distorting medium used for the presentation of a rotated field was a monocular system of prisms. This system was composed of two right angle prisms, one which could be rotated, thereby rotating the field, and the other fixed, thereby correcting the left-right reversal of the image (see Fig. 1). Each prism was placed in a brass tube 40 mm. in diameter and the tubes were fitted one in front of the other. The open ends of the tubes were covered with brass disks with circular openings in the center. Through the openings, 19 mm.

in diameter, only the refracted light fell on *S*'s right eye. This system provided a visual field of 15° visual angle. The amount of prism rotation could be read from an indicator attached to the outside of the tube (C in Fig. 1). A brass disk of 40-mm. diameter covered *S*'s left eye. This apparatus was mounted to a circular metal frame which was adjusted to *S*'s head.

(b) Visual verticality. For the measurement of visual perception of verticality, a luminous rod (50 cm. long and 20 mm. wide) was used as the stimulus-object. The luminous rod was provided by a light box which could be rotated in *S*'s frontoparallel plane. Rotation of the rod was read in fractions of a degree directly from a protractor. This rod was so placed, 2 m. before *S*, that his gaze bisected its length.

(c) Kinesthetic verticality. A metal rod, 54 cm. long and 9 mm. in diameter was used as the stimulus-object for testing the kinesthetic perception of verticality. The lower end of the rod was located one half arm's length in front of the *S* and pivoted at the level of *S*'s waist permitting rotation in the frontoparallel plane. Rotation of the rod was read directly from a protractor.

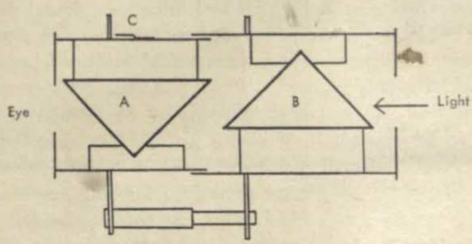


FIG. 1. DISTORTING PRISM USED IN EXPERIMENT I (A) Rotatable prism in the position of no field rotation; (B) Fixed prism; (C) Prism rotation indicator.

(d) Goggles. The goggles permitted the exposure, as *E* desired, of either eye independently. For measurement of visual verticality either the left or right eye could be exposed; for the kinesthetic verticality, both could be covered.

(2) Procedure. As schematically presented in Table I, this experiment consisted of four tests of visual and kinesthetic verticality. All tests were conducted in a dark room as follows.

(a) Visual test. *S* was seated and his head held firmly in place by a head-rest. Each eye was tested separately. *S*'s task was to inform the *E* the direction in which the lighted rod should be rotated until it appeared vertical. At the beginning of each trial, the rod was started either at 10° left or 10° right from objective vertical. Two measurements were taken for each starting-position of the rod. After one adjustment *S* was asked to close his eyes until the rod was set at the starting position for the next trial. The order of testing for eyes and starting-positions of the rod was systematically varied. Thus a total of four measures was obtained for both the left and right eye at each test.

(b) Kinesthetic test. After the goggles were placed on *S*, he was instructed to touch the rod with the finger tips of both hands. His task was to adjust the rod to

TABLE I

GENERAL PROCEDURE

Sequence of Steps

1	2	3	4	5	6	7
Pre-test	Adaptation (15 min.)	Post-test I	Readaptation (3 min.)	Post-test II	Readaptation (5 min.)	Post-test III
No prism	with prisms (All *S*s walking or seated in a light room.)	without prisms	without prisms (Exposed Group, walking or seated in a light room; Unexposed Group, seated in a dark room.)	without prisms	without prisms (Exposed Group, walking or seated in a light room; Unexposed Group, seated in a dark room.)	without prisms

apparent vertical. Two starting positions, 10° to the left and 10° to the right of objective vertical, were used. One measurement was taken for each starting-position.

(c) Measures employed. The position in which the rod was perceived, either visually or kinesthetically, as vertical (apparent vertical) was measured in degrees of deviation from the objective vertical (plumb line). Arbitrarily, an angular position clockwise of objective vertical is designated by positive numbers (plus sign) and counter-clockwise by negative numbers (minus sign). The measure used to represent *S*'s performance is the arithmetic mean of the adjustments for each test.

Results. The effect of figural adaptation in the right eye are presented in Table II. Mean positions of apparent vertical are 1.4° and 3.1° clockwise of objective vertical at Pretest and Posttest I respectively.[3] Following exposure to a titled field for 15 min. apparent vertical shifted 1.7° clockwise. This is statistically significant (Between Tests in Table II). This figural adaptation for the right eye is in general agreement with previous experiments.

The results concerning figural adaptation in the left eye (interocular

TABLE II

VERTICALITY FOR THE RIGHT EYE EXPOSED TO THE DEFLECTION: ADAPTATION

Source of variation	*df.*	Mean square	*F*
Between individuals	19	3.15	5.26*
Between tests	1	26.57	44.39*
Residual	19	0.60	—
Total	39	2.51	—

Means: Pretest, +1.4°; Posttest I, +3.4°.

* Significant at the 1% level.

[3] The physical position of the apparent vertical for the Pretest was rotated to the right. Systematic errors of this sort under control conditions have been interpreted in sensory tonic theory as an expression of habitual equilibrium of the organism. Wapner and Werner, 1957, *op. cit.*, p. 62.

TABLE III

VERTICALITY FOR THE LEFT EYE UNEXPOSED TO THE DEFLECTION: ADAPTATION

Source of variation	*df.*	Mean square	*F*
Between individuals	19	2.73	9.82*
Between tests	1	26.90	96.71*
Residual	19	0.28	—
Total	39	2.16	—

Means: Pretest, +0.9°; Posttest I, +2.6°.

* Significant at the 1% level.

adaptation) are presented in Table III. Mean positions of apparent vertical at Pretest and Posttest I are 0.9 and 2.6° clockwise of objective vertical. Following exposure of the right eye to the tilted field the apparent vertical for the left eye shifted 1.7° clockwise. This is statistically significant. Figural adaptation occurred in the unexposed eye to the same amount and in the same direction as that for the exposed eye.

The effects of intermodal figural adaptation are presented in Table IV. Means of kinesthetic apparent vertical are 1.1° counter-clockwise and 0.5° clockwise of objective vertical at Pretest and Posttest I. Following exposure of the right eye to a tilted field, figural adaptation occurred in the kinesthetic modality. This finding is statistically significant (see 'Between Tests' in Table IV).

Results concerning the rate of readaptation in the two groups are presented in Fig. 2, 3, and 4. It can be seen in Fig. 2 that the apparent vertical for the right eye shifted counter-clockwise during readaptation period, furthermore the Exposed Group showed a greater readaptation from Posttest I to II than the Unexposed Group. From Posttest II to III, the speed of readaptation was greater in the Unexposed Group than in the Exposed Group. Thus we may state that the Exposed Group readapted faster than the Unexposed Group to the normal condition.[4] This difference in the rate is statistically significant (G × T in Table V).

TABLE IV

KINESTHETIC PERCEPTION OF VERTICALITY: ADAPTATION

Source of variation	*df.*	Mean square	*F*
Between individuals	19	3.62	3.19*
Between tests	1	24.34	21.41*
Residual	19	1.14	—
Total	39	2.95	—

Means: Pretest, −1.1°; Posttest I, +0.5°.

* Significant at the 1% level.

[4] At the Pretest the physical positions of the apparent vertical for the two groups differ. This difference may be due to the use of independent groups. Since the same amount of figural adaptation was obtained for the two groups, we assume the difference at the control condition did not affect the rate of readaptation (see footnote 3).

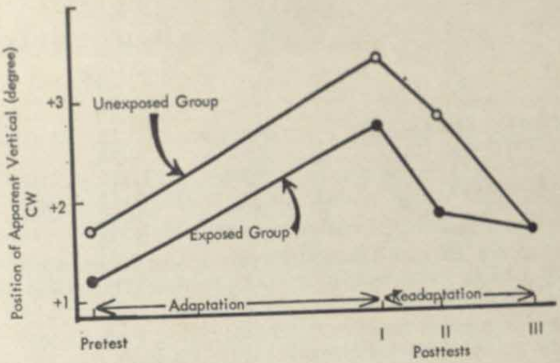


FIG. 2 SHIFT OF APPARENT VERTICAL FOR THE RIGHT EYE EXPOSED TO THE DEFLECTION

Fig. 3 presents the shift of apparent vertical for the left eye. As can be seen in Fig. 3, there is no difference in the rates of readaptation in the Exposed and Unexposed Groups. The amount of shift during the readaptation period was 1.1° clockwise in both groups (Table VI).

Fig. 4 shows the shift of kinesthetic verticality in the two groups. The Exposed Group showed complete readaptation at Posttest II. The unexposed group on the other hand showed only a slight readaptation at Posttest III. Although this difference is statistically not significant (G × T in Table VII), the mean change of the Exposed Group indicates faster readaptation than in the Unexposed Group.

In general, faster readaptation in the Exposed Group than in Unexposed

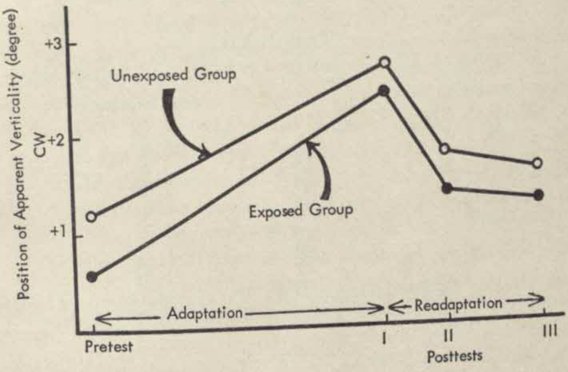


FIG. 3. SHIFT OF APPARENT VERTICAL FOR THE LEFT EYE UNEXPOSED TO THE DEFLECTION

TABLE V

VERTICALITY FOR THE RIGHT EYE EXPOSED TO THE DEFLECTION: READAPTATION

Source of variation	*df.*	Mean square	*F*
Between individuals (I):	19	5.33	—
Between groups (G)	1	5.10	<1.00[1]
Between Is within G	18	5.34	—
Within individuals:	40	0.86	—
Between tests (T)	2	11.05	39.21[2]*
Interaction G×T	2	1.15	4.08[2]*
Pooled Is×T	36	0.28	—†
Total	59	2.30	—

Means: Posttest I, Exposed Group +2.7°; Unexposed Group +3.4°; Posttest II, Exposed Group +1.8°; Unexposed Group +2.8°; Posttest III, Exposed Group +1.6°; Unexposed Group +1.6°.

[1] Tested against Between Is within G; [2] Tested against Pooled Is×T.
* Significant at the 1% level, † Significant at the 5% level.

TABLE VI

VERTICALITY FOR THE LEFT EYE UNEXPOSED TO THE DEFLECTION: READAPTATION

Source of variation	*df.*	Mean square	*F*
Between individuals (I)	19	4.32	—
Between groups (G)	1	1.87	<1.00[1]
Between Is within G	18	4.45	—
Within individuals	40	0.59	—
Between tests (T)	2	7.44	31.02[2]*
Interaction G×T	2	0.01	<1.00[2]
Pooled Is×T	36	0.24	—
Total	59	1.79	—

Means: Posttest I, Exposed Group +2.4°; Unexposed Group +2.7°; Posttest II, Exposed Group +1.4°; Unexposed Group +1.8°; Posttest III, Exposed Group +1.3°, Unexposed Group +1.6°.

[1] Tested against Between Is within G; [2] Tested against Pooled Is×T.
* Significant at the 1% level.

TABLE VII

KINESTHETIC PERCEPTION OF VERTICALITY: READAPTATION

Source of variation	*df.*	Mean square	*F*
Between individuals (I)	19	7.49	—
Between groups (G)	1	48.60	9.34[1]*
Between Is within G	18	5.20	—
Within individuals	40	1.49	—
Between tests (T)	2	4.92	3.92[2]†
Interaction G×T	2	2.27	1.81[2]
Pooled Is×T	36	1.26	—
Total	59	3.42	—

Means: Posttest I, Exposed Group 0.0°, Unexposed Group +1.0°; Posttest II, Exposed Group −1.5°, Unesxpoed Group +0.9°; Posttest III, Exposed Group −1.4°, Unexposed Group +0.5°.

[1] Tested against Between Is within G; [2] Tested against Pooled Is×T.
* Significant at the 1% level; † Significant at the 5% level.

Group was found with respect to apparent vertical for the right eye and for the kinesthetic modality. No difference was found in the rate with respect to the left eye.

Experiment II: Adaptation to a Curved Visual Field

(1) Apparatus. (a) Distorting prism. A monocular prism system composed of two 15° prisms mounted to a frame, side by side, was used as the distorting medium. For this experiment the prisms were set with their apexes together, oriented to the left. The amount of deflection used was equal to that of a 30° prism, *i.e.* 30 prism dioptors. The frame, containing the two prisms, was fitted to the right eye of a pair of goggles. The left eye of the goggles was covered. When these goggles were first

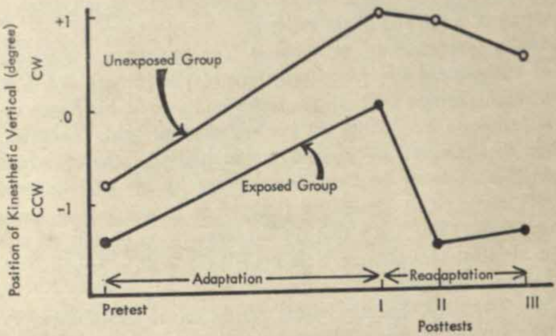


Fig. 4. Shift of Kinesthetic Apparent Vertical

worn by *S*, all vertical lines in the field appeared convexed to the right. The apparent curvature of the vertical lines was 30-mm. displacement to the right at the center of a 50 cm. vertical line. Horizontal lines in the field remained almost straight. There were other deflections in the field however. The main deflections were a lateral displacement and prismatic dispersion along the edge of vertical lines. As these deflections were identical for all *S*s, they presented no serious obstacle to the aim of the experiment. The size of the curved field was 30° of visual angle.

(b) Visual perception of straightness. A series of luminous lines with various amounts of curvature was used for the measurement of apparent straightness of lines. The luminous lines were provided by the same light-box used in Experiment I. The light box was fixed in a vertical orientation for this experiment. The luminous lines, 5 mm. × 50 cm. in size, were one straight line and curves from circles 40, 30, and 20 m. in diameter. The same curve could be presented as a line either convexed to the left or to the right. Thus seven stimulus-lines were used. This light-box was located 2 m. away from the *S* and the length of the line was bisected by his line of regard.

(c) Kinesthetic perception of straightness. The stimulus-object used for a test of kinesthetic perception of straightness was a steel bar, 1.5 mm. thick, 6 mm. wide,

and 58 cm. long. This bar was objectively straight and was vertically fixed to a stand by its ends. It was presented in front of the blindfolded *S* sitting in a chair. His shoulder bisected the length of the bar.

(d) Goggles. The same goggles employed in the Experiment I were again used for blindfolding and monocular exposures.

(2) Procedure. As schematically presented in Table I, this experiment consisted of four tests of visual and kinesthetic apparent straightness. All tests were conducted in a dark room as follows.

(a) Visual perception of straightness. *S* was seated and blindfolded. Each eye was tested separately. Stimulus-lines were presented one at a time in random order. *S*'s task was to inform *E* whether the line presented appeared straight, convexed to the right or to the left. Seven judgments of apparent straightness, one for each stimulus-line, were obtained for both eyes. When *S*'s judgment was uncertain, the line was presented once again at the end of the series.

(b) Kinesthetic perception of straightness. The apparatus for testing kinesthetic perception of straightness was placed in front of the blindfolded *S*, who touched the bar with the finger tips of both his hands and moved his hands up and down. His task was to inform *E* whether the bar appeared straight, convexed to the right or to the left. The instructions given *S* emphasized that the curavture of the bar was so slight that he must be careful and judge by his first impression. As the apparatus was covered with a piece of cloth when it was not in use, the *S* never saw the stimulus-object nor knew that the bar was always objectively straight. Three judgments of apparent straightness were made by *S* with a pause of about 10 sec. between successive judgments.

(c) Measures employed. For the visual perception of straightness, Point of Subjective Straightness *(PSS)* was determined by means of the method of constant stimuli. Arbitrary numbers were used: '20' for the curve of the greatest convexity to the right among the stimuli used, '50' for the straight line, and '80' for the curve of the greatest convexity to the left. In the tables presented below, numbers less than 50 designate *PSS* convex to the right, numbers greater than 50 designate *PSS* convex to the left, and 50 a straight line.

To determine the kinesthetic perception of straightness, arbitrary numbers were used: '5' for the judgment of convex to the left, '10' for that of straight, and '15' for convex to the right. In Table X, mean kinesthetic straightness of less than 10 indicates more judgment of convex to the left, and number greater than 10 indicates more judgment convex to the right occurred to an objectively straight rod.

Results. Results concerning figural adaptation in the right eye are presented in Table VIII. The *PSS* at Pre-test is 52.8, at Post-test I is 46.8. Thus following exposure to a curved field, the *PSS* shifted to a line with more convexity to the right. The amount of shift is statistically significant and in agreement with previous experiments concerning figural adaptation to a curved field.

Results concerning interocular adaptation are presented in Table IX. The *PSS* for the unexposed left eye shifted from a straight line to a line

TABLE VIII

Perception of Straightness for the Right Eye Exposed to the Deflection: Adaptation

Source of variation	df.	Mean square	F
Between individuals	19	66.97	2.25†
Between tests	1	360,00	12.10*
Residual	19	29.74	—
Total	39	56.35	—

Means: Pretest, 52.8, Posttest I, 46.8.

* Significant at the 1% level; † Significant at the 5% level.

TABLE IX

Perception of Straightness for the Left Eye Unexposed to the Deflection: Adaptation

Source of variation	df.	Mean square	F
Between individuals	19	57.75	2.21†
Between tests	1	390.63	14.94*
Residual	19	26.15	—
Total	39	50.88	—

Means: Pretest, 50.0; Posttest I, 43.8.

* Significant at the 1% level; † Significant at the 5% level.

TABLE X

Kinesthetic Perception of Straightness: Adaptation

Source of variation	df.	Mean square	F
Between individuals	19	90.66	2.02
Between tests	1	22.30	<1.00
Residual	19	44.87	—
Total	39	66.60	—

Means: Pretest, 10.2; Posttest, 9.7.

convexed to the right. This amount of shift is statistically significant. Thus interocular adaptation occurred with respect to perception of straightness.

Results concerning kinesthetic, apparent straightness are presented in Table X. Following exposure to a curved visual field, the means of kinesthetic apparent straightness shifted from 10.2 to 9.7. Although this shift is in the expected direction, *i.e.* number of judgment convex to the left increased, it is statistically not significant.[5] The intermodal figural adaptation was very slight with respect to perception of straightness.

Results concerning the rate of readaptation in the two groups are presented in Figs. 5 and 6. It can be seen in Fig. 5 that the Exposed Group showed a complete readaptation at Posttest II while the Unexposed Group

[5] Since the kinesthetic perception of straightness was tested with respect to the perception of an objectivity straight rod, an increase of the judgment "convex to the left" indicates a shift of the apparent straightness to a rod convexed to the right.

showed a further adaptation at Posttest II and a slight readaptation at Posttest III. These differences in the rate of readaptation are statistically significant. (G × T in Table XI).

Fig. 6 presents the shift of *PSS* for the left eye. It can be seen here that at both Posttests II and III the Exposed Group showed a greater readaptation than the Unexposed Group. The difference, however, is not significant (G × T in Table XII).

Since no significant figural adaptation was obtained for the kinesthetic perception of straightness, no analysis was made concerning the speed of

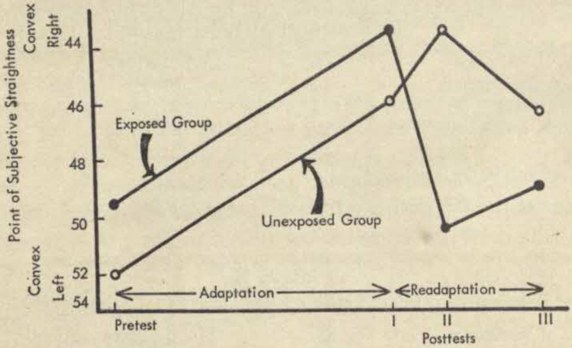


FIG. 5. SHIFT OF APPARENT STRAIGHTNESS FOR THE RIGHT EYE EXPOSED TO THE DEFLECTION

A number smaller than 50 indicates a line convex to the right; a number greater than 50 indicates a line convex to the left; 50 represents a straight line.

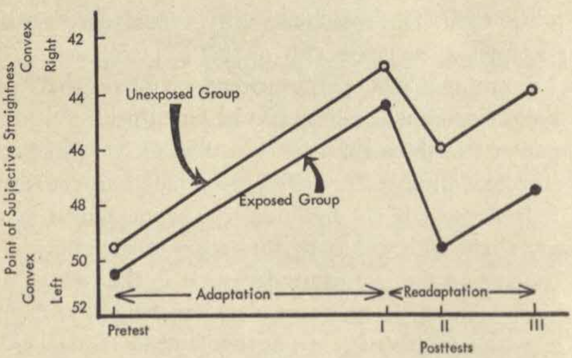


FIG. 6. SHIFT OF APPARENT STRAIGHTNESS FOR THE LEFT EYE UNEXPOSED TO THE DEFLECTION

A number smaller than 50 indicates a line convex to the right; a number greater than 50 indicates a line convex to the left; 50 represents a straight line.

TABLE XI

Perception of Straightness for the Right Eye Exposed to the Deflection: Readaption

Source of variation	df.	Mean square	F
Between individuals (I)	19	163.95	—
Between groups (G)	1	81.67	<1.00[1]
Between Is within G	18	168.52	—
Within individuals	40	46.02	—
Between tests (T)	2	48.75	1.31[2]
Interaction G×G	2	202.50	5.45[2]*
Pooled Is×T	36	37.18	—
Total	59	84.15	—

Means: Posttest I, Exposed Group 45.5, Unexposed Group 48.0; Posttest II, Exposed Group 52.5, Unexposed Group 45.5; Posttest III, Exposed Group 51.0 Unexposed Group 48.5.

[1] Tested against Between Is within G; [2] Tested against Pooled Is×T.
* Significant at the 1% level.

TABLE XII

Perception of Straightness for the Left Eye Unexposed to the Deflection: Readaptation

Source of Variation	df.	Mean square	F
Between individuals (I)	19	82.87	—
Between groups	1	120.42	1.49[1]
Between Is within G	18	80.79	—
Within individuals	40	24.17	—
Between tests (T)	2	80.00	3.63[2]*
Interaction G×T	2	6.66	<1.00[2]
Pooled Is×T	36	22.04	—
Total	59	43.07	—

Means: Posttest I, Exposed Group 44.5, Unexposed Group 43.0; Posttest II, Exposed Group 49.5, Unexposed Group 46.0; Posttest III, Exposed Group 47.5, Unexposed Group 44.0.

[1] Tested against Between Is within G. [2] Tested against Pooled Is×T.
* Significant at the 5% level.

readaptation in the two groups. In sum, it was found that the Exposed Group readapted faster than the Unexposed Group with respect to perception of straightness for both right and left eyes.

Summary and Conclusion

In the experiments reported here, figural adaptation to fields distorted by prisms was investigated. This process was studied in terms of the after-effects of adaptation to two types of distorted field, one tilted and the other curved, and readaptation back to the normal visual field. Furthermore intermodal and interocular transfer of the after-effects and readaptation was also studied. The findings are summarized as follows: In Experiment I, it was found that following exposure of the right eye to a visual

field tilted in clockwise direction, the apparent vertical for the right eye, for the unexposed left eye, and for the kinesthetic modality shifted clockwise. In Experiment II, following exposure of the right eye to a field in which all vertical lines were convexed to the right, the apparent straightness of a line for the right eye and for the unexposed left eye shifted to a line physically convexed to the right. Although not significant, the kinesthetic perception of straightness changed in the same direction as visual perception of straightness. With respect to the rate of readaptation it was found in both experiments that, in general, the group exposed to the normal visual field and allowed to move during the readaptation period showed a faster rate of readaptation in comparison to the group sitting immobile in the dark room.

The basic process of figural adaptation has been discussed as follows: assuming that perceptual experience is a function of the relation between the organismic state and proximal stimuli and that significant perceptual experiences such as verticality and straightness correspond to a stable *s/o* relationship; figural adaptation then is viewed as entailing perceptual changes based on a new relationship of stability between the deflected proximal stimuli issuing from a plumb line or physically straight line. This relationship comes about through the change of the organismic state such that this change corresponds to the shift in proximal stimulation.

The finding of interocular and intermodal figural adaptation to a tilted field and a curved field are consonant with the basic assumption underlying this interpretation of figural adaptation. Interocular and intermodal figural adaptations are at variance with the view that adaptation involves simply a local change in the sense organ directly stimulated; rather they must be attributed to a more central change in the organismic state relevant to the experience of verticality and straightness in *both* visual and kinesthetic modalities.

The second finding of this investigation concerns the effects of mobility of the *S*s and exposure to visual and tactual-kinesthetic stimuli upon the rate of readaptation. These findings though in the direction of expectations, require further experimentation before conclusions as to their theoretical significance can be drawn.

THE PERCEIVED DURATION OF AUDITORY AND VISUAL INTERVALS: CROSS-MODAL COMPARISON AND INTERACTION

By ISAAC BEHAR, Army Medical Research Laboratory, and WILLIAM BEVAN, Kansas State University

Recent studies have indicated that judgments of perceptual magnitude are related to internal norms derived from previous stimulation.[1] A central problem associated with the prediction of such judgments involves determining the criteria which identify inputs that are relevant to the derivation of this norm. One facet of this problem concerns modality of stimulation as a significant variable, particularly when the stimulus-sequence involves the presentation of signals to more than a single modality. A logical prerequisite to the investigation of such a relationship involves selecting for study a dimension which is either common to several modalities or, at least, exists for one modality and has an analogue in another. Temporal duration, which meets this requirement, was chosen as the stimulus-dimension in the present case. The practical problem of stimulus-control was handled by selecting vision and audition as the input modalities.

A common procedure in the study of the judgment is to determine the effect of a stimulus differing in magnitude from the rest of the stimulus-series.[2] The present problem may be formulated more explicitly in terms of the relative effectiveness of anchor-stimuli presented to the same modality and to a modality different from that of the stimulus-series. This, in turn, involves comparing anchor-effects in judgments of temporal duration for at least two modalities. It also involves investigating the possibility of heteromodal anchor-effects and a comparison, if they are demonstrated, with intramodal effects. In all, five interrelated experiments were performed.

---

* Received for publication January 15, 1960. This work was done at Emory University.

[1] Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29; William Bevan and C. L. Darby, Patterns of experience and the constancy of an indifference point for perceived weight, this JOURNAL, 68, 1955, 575-584.

[2] Spaulding Rogers, The anchoring of absolute judgments, *Arch. Psychol.*, 37, 1941 (No. 261), 1-42; W. A. Hunt, Anchoring effects in judgment, this JOURNAL, 54, 1941, 395-403. For a good recent summary, see J. P. Guilford, *Psychometric Methods*, 2nd ed., 1954, 312-314.

Experiment I: Anchor-Effects for Duration

The aim of the first experiment was to look for an anchor-effect in judgments of temporal duration. Although Postman and Miller had previously demonstrated an effect for auditory durations, their series was limited to the range, 0.25–1 sec., and correspondingly brief anchors were used.[3] The present design called for longer series-durations, wider ranges, and contrasting anchors.

Subjects. The Ss were 24 undergraduate students of psychology; 17 men and 7 women.

Procedure. The method of single simuli was employed in all experiments. The apparatus for producing visual signals consisted of a neon glow-lamp (G. E. NE51) mounted in a black wooden screen (18 × 24 in.) about 2 ft. in front of *S*. The intensity of the lamp was regulated with a Variac. Auditory signals were 'white noise' fed through an amplifer (Heathkit A-7) into earphones (PDR-8). Stimulus-durations were controlled by a Hunter timer (IIIC). All switches were silent, and possible cues from the setting of the timer were eliminated by the use of heavily padded earphones and an ambient masking noise.

The stimulus-series for the first experiment consisted of visual durations of 1, 2, 3, 4, and 5 sec., presented in a predetermined random order, except that: (a) no duration was repeated in succession; and (b) each duration appeared 5 times in each successive block of 25 trials. An anchor below the range of the series (0.2 sec.) and one above the series (9 sec.) were used. Each *S* judged the stimulus-series under three conditions: (a) with no anchor; (b) with the short anchor; and (c) with the long anchor. The order in which the *S*s encountered each condition was completely counterbalanced. The series was judged 15 times in each condition, making a total of 225 judgments for each *S*. Anchor-duration, when presented, were interspersed with the series as every fourth stimulus for a total of 25 presentations each. They were given no special designation by *E*.[4] An 11-category scale, *very-very-very-short* through *medium* to *very-very-very-long*, was used. The interstimulus-interval was about 6 sec. and periods of 2 min. intervened between conditions. The *S*s were not informed of the structure of the series for any condition.

At the beginning of the test-session, *S* was instructed to express his judgments verbally in terms of the 11-category scale, forming these responses relative to the durations he received, rather than in terms of any impression he might have had prior to testing. He was asked to refrain from counting or tapping as a means of estimating the intervals.

Results. The median judgment for each *S* for each series-duration was determined for each condition after converting the descriptive categories into a numerical scale by assigning the values 1–11 to the ordered categories. Next, the mean judgment for the 24 *S*s for each of the 5 durations

---

[3] L. J. Postman and G. A. Miller, Anchoring of temporal judgments, this Journal, 58, 1945, 43-53.

[4] What is here called an *anchor*, Helson refers to as a predominant-stimulus member of the series.

in each of the 3 conditions was calculated. A 3 × 5 × 24 (treatments × treatments × *Ss*) analysis of variance was performed on these data.[5] All main effects were significant: among *Ss*, among durations, and among conditions; the former two at the 0.1% level ($F_{23/46} = 11.83$; $F_{4/92} =$ 21[illegible].88) and the latter at the 5% level ($F_{2/46} = 3.46$). The data are presented graphically in Fig. 1.[6]

Because the order of conditions was completely counterabalanced, it was possible to extract from the *SS* for *Ss* a component associated with se-

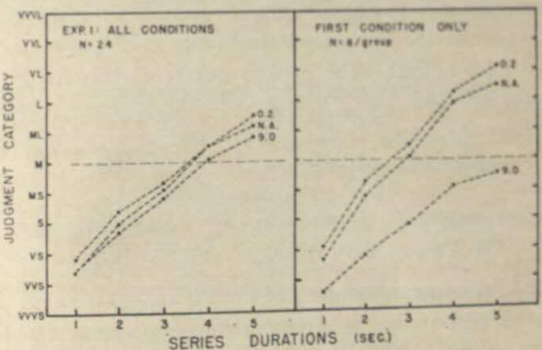


FIG. 1. INTRAMODAL ANCHOR-EFFECT FOR VISION WITH ANCHORS ABOVE AND BELOW SERIES

(The left panel shows the data of Experiment I across all orders; the right panel shows the data for the condition presented first.)

quences of conditions, and, from the *SS* for *Ss* × conditions, a sequence × conditions interaction. Both were significant ($F_{5/18} = 4.49$, $P < 0.01$; $F_{10/36} = 2.25$, $P < 0.05$). These significant *F*-ratios indicate a perseveration of the effects of one condition to the next. This perseveration might be expected to reduce the observed anchor-effect when, as in the present analysis, the latter results from the averaging of judgments for each condition regardless of serial order. Such an attenuation should contrast sharply with the condition presented first, where there can be no perseveration. Accordingly, the mean judgments of the 8 *Ss* in each condition for

[5] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1950, 237-238.

[6] The mean responses for the several conditions of each of the remaining experiments, as in this case, are presented graphically. Since the series serves primarily to demonstrate that the mechanism of pooling is operative across modalities, rather than to delineate a quantitative version of the pooling principle, precise tabulation of these data has been omitted in the interest of space. Such data are available upon request.

the first-presented condition were determined, and evaluated in a mixed-design analysis of variance.[7] In this case, the differences among conditions are very large, confirming the present hypothesis. Furthermore, the effect of the long anchor exceeds that of the short by more than four times. The difference among conditions presented initially is significant at the 0.1% level ($F_{2/21} = 10.24$).

The analysis of both the over-all results and those for the first-presented condition alone confirm the occurrence of anchoring effects with longer durations than have heretofore been demonstrated. The perseverative effect, noted also by earlier workers, strongly indicates the necessity of using independent groups for the several conditions to obtain an accurate estimate of the anchor-effect.[8] The marked difference between the long and short anchor-effects indicates a limitation in the range of possible anchor-durations below the series. It was thus decided to use only longer anchor durations in subsequent experiments.

Experiment II: Comparison of Anchoring with Visual and Auditory Durations

The second experiment was directed toward a comparison of the relative effectiveness of anchors in the two modalities of vision and audition. Separate anchor and no-anchor groups were used in both modalities to avoid inaccurate estimates of the anchor-shift associated with perseverative effects.

Subjects. The *S*s were 120 undergraduate students of psychology; 73 men and 47 women.

Procedure. For convenience of discussion, this study is described in terms of two sets of data: one involving the judgment of light-filled intervals (Experiment II-L), the other the judgment of noise-filled intervals (Experiment II-N). In Experiment II-L, there were 75 *S*s divided randomly (except to equate sex-ratios) into five groups. The control group judged the series alone, the others, the series plus a 6-, 10-, 15-, or 20-sec. anchor. Each series-duration was presented 30 times in random sequence; each anchor, as every fourth trial, 50 times. Although *S*s began judging with the first stimulus, the first 25 judgments were treated as practice and not included in computations.

Experiment II-N, involving three groups of 15 *S*s each, was a replication of the no-anchor, 10- and 20-sec. anchor-conditions of Experiment II-L, except that the durations were auditory. The perceived intensity of the noise was adjusted by heteromodal subjective-magnitude matching to the apparent intensity of the neon light.

---

[7] Lindquist, *op. cit.*, 266-305.

[8] S. R. Truman and E. G. Wever, The judgment of pitch as a function of the series, *Univer. Calif. Publ. Psychol.*, 3, 1928, 215-223; D. M. Johnson, Learning function for a change in a scale of judgment, *J. exp. Psychol.*, 39, 1949, 851-860; M. E. Tresselt, The influence of the amount of practice upon the formation of a scale of judgment, *ibid.*, 37, 1947, 251-260.

The value used was the median of the equality-judgments of three experienced judges.

Results. The means for the 15 *Ss* in each group of Experiment II-L are represented in the left panel of Fig. 2; those for the groups of Experiment II-N in the right panel. Mixed-design analyses of variance were performed on the two sets of data. The obtained differences in judgment for both visual and auditory groups are significant at the 0.1% level ($F_{4/70} = 16.86$; $F_{2/42} = 17.8$). The interactions of anchor-conditions × durations are equally reliable ($F_{16/280} = 45.4$; $F_{8/168} = 10.17$). These data reveal, as have those of earlier studies, that the extent of an anchor's influence is

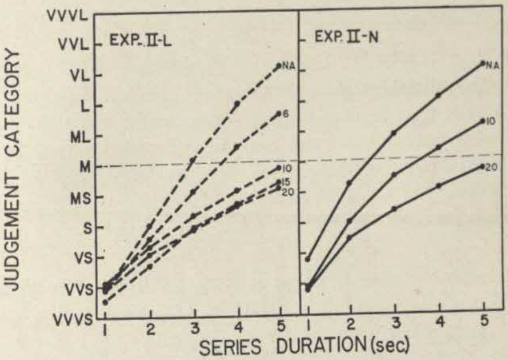


FIG. 2. INTRAMODAL ANCHOR-EFFECTS FOR VISION (LEFT PANEL) AND AUDITION (RIGHT PANEL)

proportional to its magnitude; meanwhile, our present results suggest the function to be negatively accelerated as the limit of anchor-effectiveness is approached. The anchor-effects are asymmetrical, those series durations closest to the anchor showing the greatest shift.

While the results for the two modalities are, in general, quite similar, certain differences must be noted. A comparison of the three pairs of corresponding groups indicates that the auditory durations were judged longer than the visual durations. This difference was clearly significant ($F_{1/84} = 8.26$, $P < 0.01$), but the significant interaction (modality × duration, $F_{4/336} = 3.83$, $P < 0.01$) indicates that this difference was not consistent throughout. Reversals occurred only for the longest duration in the no-anchor conditions, and the shortest duration in the 10- and 20-sec. anchor-conditions. As there is, in addition, a significant anchor-conditions × modalities × durations interaction ($F_{8/336} = 6.73$, $P < 0.001$), it is not clear

at this stage of experimentation whether the observed modalities-effect reflects a true subjective difference in the duration of auditory and visual signals, or whether it is the result of greater anchor-effects under visual than under auditory conditions.

Experiment III: Audio-Visual Difference in Time-Perception

The preceding experiment revealed that *S*s judging auditory intervals consistently used categories indicating longer durations than did those judging visual intervals. The present study was designed to check the existence of this modality-difference by eliminating the possibility of differential intermodal anchor-effects, and to obtain an estimate of its magnitude.

Subjects. The *S*s were 20 undergraduate students of psychology; 16 men and 4 women.

Procedure. Each *S* judged 75 visual and 75 auditory signals with durations of from 1–5 sec. The signals were presented in a randomly-ordered, mixed series, except that: (a) runs were limited to three successive presentations in one modality; and (b) there were five auditory and five visual presentations of each duration in successive blocks of 50 trials. Instructions were like those previously used. At the end of the test session, as a further check, eight of the *S*s were questioned concerning perceived differences in the subjective intensity of the signals in each modality.

Results. The median judgment for each duration was determined separately for the auditory and visual trials. The means of these values for the 20 *S*s are presented graphically in the extreme left panel of Fig. 3. It is clear from this figure that the data of this experiment confirm the greater (about 20%) judged duration of the auditory signals suggested by the preceding study and also reported by others.[9] Of the 8 *S*s who made judgments of the relative intensity of the signals for the two modalities, three judged them to be equal, while five judged sound to be slightly more intense. Meanwhile, inspection of the individual data revealed no systematic relationship between subjective impressions of intensity and the demonstrated intermodal difference in judged duration. An analysis of variance (treatments × treatments × *S*s) further indicates that the intermodal difference is highly significant ($F_{1/19} = 39.7$, $P < 0.001$). The interaction of modalities and durations is not significant ($F_{4/76} = 1.43$), indicating the relative over-estimation to be approximately constant, at least over the 1–5 sec. range.[10]

---

[9] Sanford Goldstone, W. K. Boardman, and W. T. Lhamon, Intersensory comparison of temporal judgments, *J. exp. Psychol.,* 57, 1959, 243-248.

[10] A more recent analysis indicates that the difference may appear from the initial judgment. (See Isaac Behar and William Bevan, Analysis of the prime psychophysical judgment, *Percept. Mot. Skills,* 10, 1960, 82.)

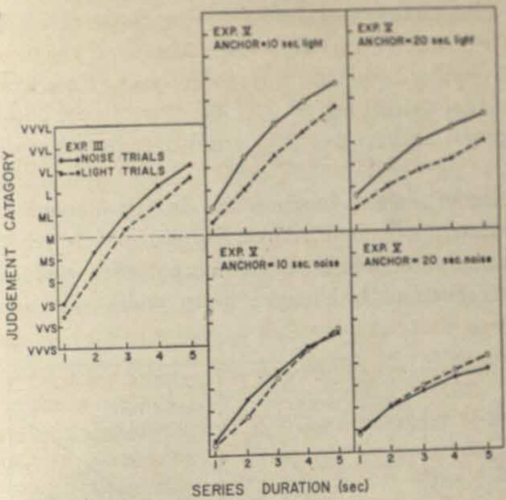


FIG. 3. CONTROL-LEVELS OF EXPERIMENT III (LEFT PANEL) AND ANCHOR-EFFECTS FOR MIXED AUDIO-VISUAL SERIES (RIGHT PANELS)

EXPERIMENT IV: HETEROMODAL ANCHOR-EFFECTS

Earlier students of anchoring have, with few exceptions, restricted their manipulation of stimuli to variations along a single input-dimension. The present experiment explores the possible effect upon judged duration of introducing anchor-stimuli to one modality and series-stimuli to another.

Subjects. The *Ss* were 60 undergraduate students of psychology; 31 men and 29 women.

Procedure. The *Ss* were assigned at random (except for equating sex-ratios) to four groups. Two groups received visual series-stimuli and two auditory series-stimuli. One group in each pair received a 10-sec. and the other a 20-sec. anchor from the alternate modality. The testing procedure was essentially the same as in Experiment II.

Results. The several group means of the median judgments are presented graphically in Fig. 4, which also includes, for reference-levels, the no-anchor means of Experiment II. Inspection clearly indicates that the heteromodal anchors, like the intramodal anchors of earlier experiments, are capable of inducing shifts in judged series-durations. Mixed-design analyses of variance were performed separately on the data of the three visual-series groups (left panel of Fig. 4) and on the data of the three auditory-series groups (right panel). For both modalities, the differences among the no-anchor condition and the two heteromodal anchor-conditions

are highly significant (visual series: $F_{2/42} = 9.35$, $P < 0.001$; auditory series: $F_{2/42} = 12.96$, $P < 0.001$). The anchor interaction (conditions × durations) is equally significant in both analyses ($F_{8/168} = 32.10$; $F_{8/168} = 10.15$), indicating that the effect of the heteromodal anchors, like that of the intramodal anchors, is asymmetrical over the range of series-durations.

The influence of the visual anchors was directly proportional to anchor-magnitude, while an opposite relationship obtained for the auditory anchors. When compared with intramodal anchor effects (Fig. 2), the magnitude of the effects of both 10-sec. heteromodal anchors is seen to be

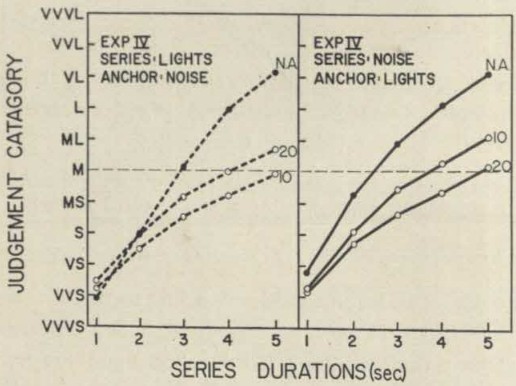


FIG. 4. HETEROMODAL ANCHOR-EFFECTS FOR VISUAL SERIES (LEFT PANEL) AND AUDITORY SERIES (RIGHT PANEL)

about equal to that obtained with corresponding intramodal durations. By contrast, the effects of the 20-sec. heteromodal anchors are relatively reduced (the heteromodal effects for visual and auditory anchors are only 67% and 22% of their intramodal effects, respectively).

Anchor-effectiveness is thus seen to be influenced both by anchor-magnitude and anchor-modality. For anchors close to the series-magnitudes, heteromodal anchors are equivalent to intermodal anchors. With anchor-magnitudes further removed from the series, however, there is a loss in anchor-effectiveness which is greater for auditory than for visual heteromodal anchors.

EXPERIMENT V: ANCHOR-EFFECTS WITH MIXED AUDIO-VISUAL SERIES

The aim of the final series of experiments was to determine the effects of anchors of a single modality on a mixed audio-visual series.

Subjects. The *Ss* were 40 undergraduate students of psychology; 22 men and 18 women.

Procedure. Each *S* was assigned at random, equating sex-ratios, to one of four anchor-condition groups. All groups judged the mixed series of visual and auditory durations used in Experiment III, but with an anchor-duration interpolated into the series as in Experiments II and IV. A visual anchor of either 10 or 20 sec. was used with two groups; auditory anchors of the same durations were used with the remaining two. The procedure was identical with that used in Experiment III, except as the inclusion of anchor-stimuli required modification.

Results. For each *S*, the median judgment for each duration was determined separately for the the trials in each modality, as in Experiment III. The group-means of these median judgments for each anchor-condition are represented in Fig. 3, along with the control-data from Experiment III. The mixed-design analysis of variance performed on these combined results indicates that, of the four main variables, the differences among anchor-durations ($F_{2/54} = 36.7$, $P < 0.001$), series-durations ($F_{4/216} = 653.6$, $P < 0.001$), and series-modalities ($F_{1/54} = 72.8$, $P < 0.001$) are highly reliable, while that between anchor-modalities is not ($F_{1/54} = 1.76$). This last effect implies a comparable over-all influence of the visual and auditory anchors.

Examination of Fig. 3 reveals that the anchors of both modalities had a greater intramodal than heteromodal effect. In comparison to the modality-difference obtained in Experiment III, visual anchors produced a greater difference in the judged duration of visual and auditory stimuli, while a reduced modality-difference results with auditory anchors. Indeed, the 20-sec. auditory anchor produced a reversal in the direction of the modality-difference. This differential effect of anchor-modality on intramodal and heteromodal durations is confirmed by the significant series modality-anchor modality interaction ($F_{1/54} = 38.9, P < 0.001$).

Finally, the intramodal *vs*. heteromodal anchor effects described are asymmetrical, *i.e.* the differential effect of anchor-modality is greater for those series-inputs closest to the anchor, as established by the significant series-modality × anchor-modality × series-duration interaction ($F_{4/216} = 3.17$, $P < 0.05$). The simple asymmetrical effect of the anchors on the judgments of the members of the series found in Experiments II and IV are again confirmed in this study; the interaction of series-duration × anchor-duration is highly significant ($F_{8/216} = 21.9, P < 0.001$).

Discussion

The results presented, when viewed as a whole, are essentially uncomplicated and require little special comment. They confirm and extend

earlier findings of anchor-effects for temporal duration. By demonstrating that judgments of duration have properties in common with judgments of material dimensions, they point up the conceptual nature of all dimensions. While anchors of duration longer than the series-stimuli were more effective than those of shorter duration, this result may simply reflect, in part at least, the limited range below the present series in which the anchor could be manipulated.

The demonstration of heteromodal anchor-effects indicates that modality is not a limiting factor in the identification of relevant input for pooling. It also supports the view that sensory data, so long attributed to the operation of peripheral mechanisms, reflect a complex judgmental process, largely central in character. Consistent with this interpretation is the fact that the anchor-effects in the two modalities studied, vision and hearing, were highly similar. The greater effectiveness of intramodal anchors, especially the more distant ones, suggests that modality is also a determining condition of degree of relevance. The greater effectiveness of visual heteromodal anchors suggests that they are subjectively closer to the series than their auditory counterparts. Finally, the fact that auditory intervals were consistently judged longer than visual intervals of the same duration calls attention to the contribution of peripheral variables and indicates that they must not be ignored in accounting for psychophysical judgments.

Summary

The main purpose of the present study was to investigate intermodal relationships in judgment by comparing the shift of subjective scales for duration with intramodal and heteromodal anchors. Five experiments, with 264 *Ss*, were conducted, in each of which the method of single stimuli was used, with a standard series of five arithmetically-spaced intervals of 1–5 sec. Two modalities, vision and audition, were compared. Anchor-durations usually were 10 and 20 sec.

The results of the series justify the following conclusions: (a) Intramodal anchor-effects are readily elicited, and are similar for both modalities, vision and audition. (b) Auditory intervals are judged to be about 20% longer than visual intervals. (c) Heteromodal anchor-effects also can be induced, but differ in magnitude from comparable intramodal effects.

ACCURACY AND CONSISTENCY OF TIME-ESTIMATION BY FOUR METHODS OF REPRODUCTION

By D. G. Doehring, Central Institute for the Deaf

The problem of determining how well an individual can estimate the passage of time has been attacked in a number of different ways. The conflicting results obtained from a large number of investigations over the past 75 years do not support the notion that individuals make use of a stable 'internal clock' to judge the passage of time. Rather, the individual's performance appears to be sensitive to the exact operations that are used to elicit his estimates.[1] This suggests there is need to specify more adequately the variables that influence these judgments.

The present study varied in several ways the common procedure for judging time; namely, the method of reproduction of empty intervals. With this procedure, *E* presents the standard, an 'empty' interval of time bounded by instantaneous sounds, and *S* is required to reproduce the standard by two taps on a response-key. Where a number of trials are given with each standard, the precision of judgment is defined in terms of both the central tendency and the variability of *S*'s estimates. In the following discussion, the central tendency of estimates will be denoted by the term *accuracy* and the variability of estimates will be denoted by the term *consistency.*

When *S* is asked to reproduce the standard by tapping a telegraph key, it can be hypothesized that certain characteristics of the resulting muscular responses will influence *S*'s estimate. He cannot reproduce intervals shorter than 0.1 sec., since his maximal tapping speed is about 10 per sec. From this interval to an interval of about 1 sec., the two taps appear to constitute a continuous series of movements. For intervals of about 2 to 4 sec., the taps tend to become separate movements, but a certain amount of tension may be maintained during the interval between taps. For intervals of 4 sec. or longer, there may be unsystematic changes in tension during the period between taps. Thus, for intervals ranging from about 0.2 sec. to about 4 sec., the cues arising from muscular activity may increase the consistency

* Received for publication December 18, 1958. Supported by grants from the National Institute of Neurological Diseases and Blindness (B-243) and the National Science Foundation (G-4457). The author is indebted to Joseph Rosenstein for assistance in collecting the data reported in this study.

[1] Herbert Woodrow, Time perception, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 1224-1236.

of time-judgments, since *S* can pace his performance by maintaining a uniform rate of movement or by making use of proprioceptive cues that serve to connect the tapping responses. Estimates of intervals of 4 sec. or longer should be less consistent, because the bursts of muscular activity involved in tapping may be completely separated by extraneous activity. Woodrow, using the method of reproduction of empty intervals, obtained results that are in accordance with these predictions.[2] If it is assumed that

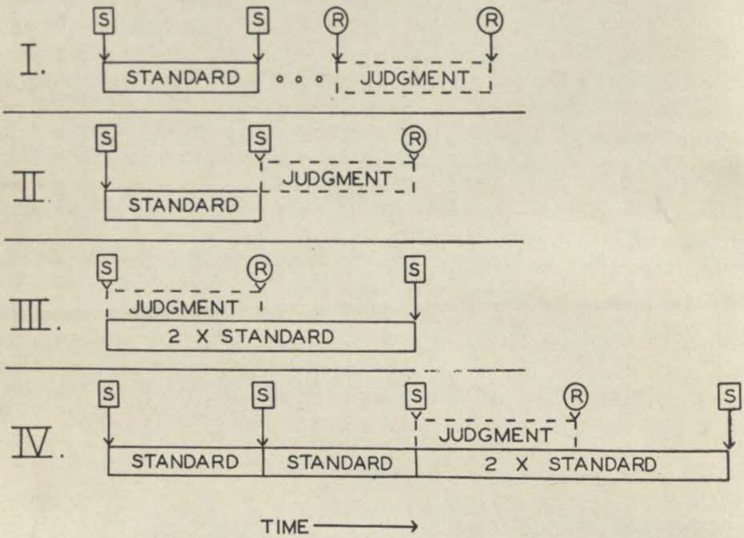


FIG. 1. METHODS OF TIME-ESTIMATION EMPLOYED
The stimuli (S) were clicks, and the response (R), the closing of a telegraph key, also produced a click.

a decrease in consistency merely increases the variability of estimates, accuracy of judgment should be unaffected by changes in the pattern of muscular responses.

METHOD

Four variations of the method of reproduction were employed in the present experiment to determine the influence of response-factors on the accuracy and consistency of time-estimation. Fig. 1 shows the four procedures that were used. Method I is the usual method of reproduction, as dissused above, where *S* estimates the duration of the standard by giving two taps on a telegraph key. Methods II, III, and IV eliminate some of the cues that might result from overt movements by re-

[2] Woodrow, The reproduction of temporal intervals, *J. exp. Psychol.*, 13, 1930, 473-499.

quiring only one tap for each estimate. With Method II, the standard is presented in the usual manner, and *S* then attempts to complete the second of two consecutive equal intervals by adding a third click to the standard. This procedure permits *S* to maintain a rate of clicks that is imposed by the two clicks of the standard. In Method III, *S* is presented with an interval that is twice the length of the standard, and is asked to bisect this interval on subsequent trials. This procedure does not provide *S* with complete information about the duration of the standard, but does give him more information about the accuracy of his estimate, since an incorrect response produces a change in both the interval preceding and the interval following the response. With Method IV, *S* is presented with three equally-spaced clicks, which constitute two consecutive presentations of the standard. These are followed by a fourth click after an interval that is twice the length of the standard. *S* is required to bisect the interval between the third and the fourth clicks. This method should be optimal for time-judgment without the aid of cues produced by movement, since it essentially combines the advantages of Methods II and III.

Intervals of 0.5, 1, 2, 4, and 8 sec. were chosen as standards for this experiment. If cues produced by movement increase the consistency of judgment in the manner described above, Method I should produce more consistent judgments than the other three methods for the three shortest intervals, and there should be no difference between Method I and the other methods for the 4-sec. and 8-sec. intervals. Several additional comparisons among methods can be made. If exact information about the standard is important for consistency of judgment, Method IV should produce the most consistent judgments and Method III should produce the least consistent judgments. If knowledge of results is more important than exact information about the standard, Methods III and IV should tend to produce more consistent judgments than Methods I and II. Accuracy of the judgments made with these methods may depend upon factors which have not yet been specified. A discussion of these factors will be reserved for a later section of this paper.

Subjects. The *S*s were eight men, including two high school students, two college students, and four college graduates.

Design. Each of the five time-intervals was judged by each of the four methods shown in Fig. 1, making a total of 20 experimental conditions. Each *S* was given 10 consecutive trials with each condition. The conditions were presented in a different random order to each *S*.

Apparatus. The auditory pulses which delimited the standard were generated by a loudspeaker fed from electronic timing circuits. The output of an audio-oscillator was led through a Schmitt trigger circuit to two decimal counting units in cascade that served as a scale-of-one-hundred frequency-divider. The output of the divider was fed through either one or three additional decimal counting units to a pulse generator, whose output was in turn fed in parallel to a loudspeaker located in *S*'s room and to one channel of a four-channel polygraph.

For the two-click standards used in Methods I, II, and III, the output of the frequency-divider was fed through one counting unit, with a pulse being generated at the beginning and the end of the count. For the four-click standard used in Method IV, the output of the frequency divider was fed through a series of three counting units, with the final pulse of one counting cycle acting as the initial pulse of the next counting cycle. The interval between pulses could be varied by changing the frequency of the audio-oscillator and by changing the number of counts that

were pre-set on the counting units. The pulse-generator could also be activated on the closing of a response-key. Thus, the standard and *S*'s estimate of the standard consisted of time-intervals that were bounded by brief auditory pulses, and a graphic record of the pulses was obtained.

The paper speed of the polygraph was adjusted in such a way that the duration of the standard was recorded as a distance of either 12 or 15 mm., thus making the precision of measurement about the same for each standard.

The auditory pulse was heard as a click. Peak amplitude of the click at *S*'s head was 67 *db, SPL* as measured by a General Radio Sound Level Meter in combination with a General Radio Impact-Noise Analyzer.

Procedure. *S* was seated in a comfortable chair in a darkened, quiet-room. The response-key was attached to a rectangular board that was placed on one arm of the chair, depending upon *S*'s preferred hand. *E* communicated with *S* through an intercommunication system from the control room in which the apparatus was located. *S* was instructed not to use any form of counting or rhythmical tapping in estimating a time-interval. The procedure appropriate to a given method was explained to *S* the first time that the method was used, and this explanation was repeated whenever necessary. Before each series of trials with a given condition, *S* was given practice, which he continued until he thought that his estimate could not be further improved. The 20 experimental conditions were presented in 2, 3, or 4 sessions, during which he was told to request a rest-period whenever he became bored or tired. The interval between trials was varied unsystematically from 1 to 6 sec. When Method I was used, *S* was allowed to respond whenever he wished after the standard had been presented. For Method III and for all intervals of 0.5 sec., a total of 12 or 13 trials was given, with only the last 10 trials being recorded.

Treatment of results. All estimates were expressed as a percentage of the standard. A mean and *SD* of 10 judgments under each of the 20 conditions was calculated for each *S*. The distributions of both the means and the *SD*s were found to be positively skewed. For purposes of statistical analysis, therefore, the individual means and *SD*s were converted to logarithms. The primary test for statistical significance was the analysis of variance for repeated measurements of the same *S*s.[3] Where a significant effect was found with this analysis, the differences among individual means were further analyzed by the Newman-Keuls sequential-range test.[4] The 5% level was accepted as the criterion for significance in all analyses.

RESULTS

Consistency. Fig. 2 shows the average *SD* for each time-interval with each method. Every point represents a mean of the unconverted *SD*s of the eight *S*s. For Method I, as had been predicted, judgments of the three shortest intervals were much more consistent than judgments of the two longest intervals. Contrary to the prediction, however, this general relation was also found for the other three methods. There were no striking dif-

---

[3] A. L. Edwards, *Experimental Design in Psychological Research,* 1950, 284-297.
[4] D. B. Duncan, Multiple range and multiple *F*-tests, *Biometrics,* 11, 1955, 1-42.

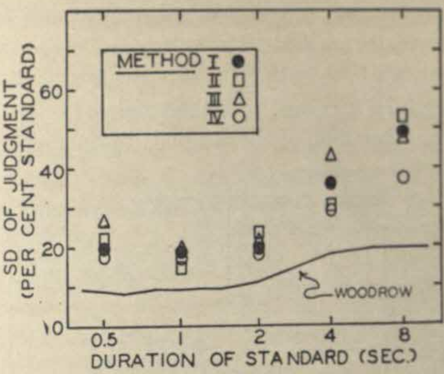


FIG. 2. CONSISTENCY OF JUDGMENTS

Each point represents the mean of 8 measures of *SD*, with the measures being based upon the 10 judgments that were made by each of the 8 *S*s. The solid line shows the results obtained by Woodrow for intervals of 0.4, 0.6, 0.8, 1.0, 1.2, 1.5, 2.0, 4.0, 6.0, and 10.0 sec.

ferences among methods either in the overall magnitude of *SD*s or in the magnitude of *SD*s as a function of duration of the standard.

Results of the analysis of variance of the *SD* of judgments, converted to logarithms, are shown in the last two columns of Table I. As might be expected from Fig. 2, the effect of duration was highly significant, and further analysis by the Newman-Keuls test revealed that judgments of the 8- and 4-sec. intervals were significantly less consistent than judgments of the three shorter intervals, and that judgments of the 8-sec. interval were significantly less consistent than judgments of the 4-sec. interval. There

TABLE I

ANALYSIS OF VARIANCE USING INDIVIDUAL MEAN JUDGMENTS AND INDIVIDUAL *SD*s

		Log Mean		Log *SD*	
Source	*df.*	Mean square	*F*	Mean square	*F*
Subjects (*S*)	7	11025		845.71	
Methods (M)	3	4111	—	913.33	4.42*
Mx*S*	21	6053		206.43	
Durations (D)	4	15527	6.95*	9815.00	24.85*
Dx*S*	28	2235		395.04	
MxD	12	3068	1.18	193.33	1.69
MxDx*S*	84	2595		114.57	
Total	159				

* Significant beyond the 5% level.

was also a significant effect of the method, and further analysis showed that judgments made by Method III were significantly less consistent than judgments made by Method IV. This was the largest difference among methods. As might be expected from the similarity of trends among the four methods, the interaction of the methods by duration was not significant.

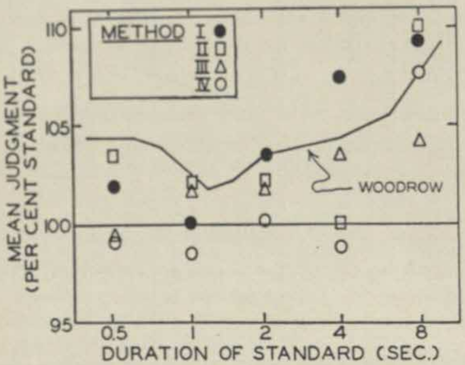


FIG. 3. ACCURACY OF JUDGMENTS

Each point represents the mean of 80 judgments, with ten judgments made by each of the eight *Ss*. The solid line shows the results obtained by Woodrow for intervals of 0.4, 0.6, 0.8, 1.0, 1.2, 1.5, 2.0, 4.0, 6.0, and 10.0 sec.

The solid line in Fig. 2 shows the results obtained by Woodrow for judgments of 10 intervals ranging from 0.4–10 sec. in duration.[5] Woodrow's *Ss* exhibited the same decrease in consistency of judgments at 4 sec. as did the *Ss* of the present experiment. The greater over-all consistency exhibited by Woodrow's *Ss* may be attributed to the fact that his *SDs* were based on a total of 100 trials for each *S* with each interval.

Accuracy. Fig. 3 shows the unconverted mean judgments obtained for each interval with each method. There appear to be no systematic differences among the four methods. Judgments of the three shortest intervals tended to approach the duration of the standard. This trend was continued for judgments of the 4-sec. interval by Methods II and IV, but judgments by Method I considerably overestimated this interval. Within each of the four methods, the largest amount of overestimation occurred for judgments of the 8-sec. interval.

[5] Woodrow, *op. cit.*, 1930, 473-499. Woodrow used a slightly different measure of variability. His measures, as shown in Fig. 2, have, therefore, been converted in such a way that they conform to the measure of variability that was used in the present experiment.

Results of the analysis of variance of log mean-judgment are shown in the third and fourth columns of Table I. The nonsignificant F-ratio for methods indicates that there were no significant differences in over-all level of judgment among methods. There was, however, a significant durative effect, and a further analysis by the Newman-Keuls test revealed that the duration of the 8-sec. standard was significantly overestimated as compared with each of the other four intervals. The interaction of durations by method was not significant, indicating that there were no significant differences in trend among the four methods.

The average accuracy of Woodrow's *S*s is shown by the solid line in Fig. 1. Their judgments were very similar to the judgments made by the eight *S*s in the present experiment.

Discussion

Consistency. The results of this experiment do not support the hypothesis that cues arising from muscular responses tend to increase the consistency of estimates of relatively short time-intervals. Judgments of the shorter intervals were found to be more consistent than judgments of the longer intervals even where the possibility of such cues had been eliminated. The similarity of Woodrow's results indicates that this trend was not peculiar to the *S*s that were tested in the present experiment. The reason for the sharp decrease in consistency of judgment at about 4 sec. thus remains undisclosed.

Judgments made with Method IV were found to be significantly more consistent than judgments made with Method III. This indicates that exact information about the standard is important, since the only difference between the two methods was a double presentation of the standard prior to bisection for Method IV. It cannot, however, be concluded that exact information about the standard is more important for consistency than knowledge of results, since Method IV involved both of these factors. The lack of a significant difference between Methods II and III suggests that these factors may operate in a similar manner in the production of consistent estimates.

Accuracy. Variations in the procedure that was used for the reproduction of empty intervals did not significantly change either the over-all accuracy of judgment or the relation between accuracy of judgment and duration of the standard. This indicates that cues arising from continuous movement or from sustained muscular tension do not bias *S*'s estimate of the standard. Once again, the results obtained in the present experiment are consistent with Woodrow's results.

Relation to other investigations. Although little attention has been paid to the consistency of time-judgments, the accuracy of time-judgment by the method of reproduction has been studied by many investigators.[6] In many cases their results appear to confirm Vierordt's Law, which states that there is a tendency to overestimate the duration of short intervals and to underestimate the duration of long intervals. Where such results have been obtained, there is an indifference-interval

[6] Woodrow, *op. cit.*, 1951, 1224-1236; The temporal indifference interval determined by the method of mean error, *J. exp. Psychol.*, 17, 1934, 167-188.

at the point where overestimation changes to underestimation. Time-estimation is obviously most accurate at this point, since the average error of estimation is zero. The exact location of the indifference-point, however, appears to be a function of the range of time-intervals that are used in a given experiment. Furthermore, some investigators have found no indifference-point and others have found a reversal of the usual trend from overestimation to underestimation.

Woodrow has conducted a number of investigations of time-judgment by the method of reproduction.[7] In the experiment that was used for comparison with the present experiment, his procedure was very similar to that used for Method I of the present experiment. Each of the eight *Ss* was given 50 consecutive trials with each interval, with the intervals presented in an ascending order from shortest to longest; this was followed by 50 consecutive trials with each interval presented in a descending order. The results of this experiment, which had been carefully planned and executed, did not confirm Vierordt's Law, as can be seen in Fig. 3. Woodrow attributed this unexpected discrepancy to differences among *Ss* in the way that they carried out instructions.[8] The similarity of the results obtained in the present experiment seems to rule out this explanation. An alternate explanation is suggested by the fact that in both experiments the trials with a given interval were presented consecutively, whereas the most common procedure for investigating time-judgment is to randomize the order of presentation of intervals from trial to trial. With the latter method, conditions are optimal for the operation of the effect of adaptation, where all estimates would tend to regress toward the middle of the range of time-intervals. This tendency should be minimized in experiments where all trials with a given interval are presented consecutively.

The results of a previous, unpublished study by the writer provide additional evidence for such an effect. Estimates of seven time-intervals were obtained by the usual method of reproduction. Eight estimates of each interval were obtained from six *Ss*. This experiment employed the same equipment that was used in the present experiment, and the procedure was the same as that used for Method I of the present experiment except that the duration of the standard was varied randomly from trial to trial. The accuracy and consistency of the resulting judgments are shown in Table II. There was a tendency to overestimate the shorter intervals and to underestimate the longer intervals, in accordance with Vierordt's Law. This trend was statistically significant. The results also differed from the results of the present experiment in terms of consistency. There was a decrease in the consistency of judgments of the 4-sec. interval, but judgments of the longer intervals became more consistent instead of less consistent. Moreover, the differences among intervals were not statistically significant.

Variation in the method of programming trials appears to produce a greater change in the pattern of estimates than variation of the response required for reproduction of the standard. It may be concluded that the method used in Woodrow's experiment and in the present experiment provides a 'true' measure of the accuracy of time-judgment, since it reduces the possibility of adaptative effects. Such a generalization is rendered unfeasible by the results of a later experiment by Woodrow, where he completely eliminated the possibility of adaptative effects by using a different group of *Ss* for each time-interval.[9] With intervals of 0.3, 0.6, 0.7, 1.2, and

[7] Woodrow, *op. cit.*, 1951, 1224-1236.
[8] Woodrow, *op. cit.*, 1934, 171.
[9] Woodrow, *op. cit.*, 1934, 167-188.

4.0 sec., the shortest interval was overestimated, the next longest interval constituted the indifference-point, and the three longest intervals were progressively underestimated. Vierordt's Law has thus been confirmed for situations which permit the maximal effect of adaptation and for situations which completely eliminate the possibility of this effect, but the law does not seem to operate in a situation that is between these extremes.

TABLE II

ACCURACY AND CONSISTENCY OF JUDGMENTS WHEN DURATION OF THE STANDARD WAS VARIED FROM TRIAL TO TRIAL

Duration of standard (in sec.)	Judgment of standard	
	Mean	*SD*
1	119.6	19.1
2	120.3	19.4
4	110.5	26.6
6	105.5	23.9
8	98.8	21.5
10	98.1	21.8
12	91.1	19.2

There appears to be no general rule for predicting the course of time-judgments with respect to duration of the standard, even when the predictions are restricted to judgments that are made by the reproduction of empty intervals. In conclusion, we can only state once again that attempts to measure an invariant faculty of time-judgment should be abandoned in favor of investigations whose purpose is to specify the variables which influence these judgments.

SUMMARY

Accuracy and consistency of time-judgment by four methods of reproduction were studied, with time-intervals of 0.5, 1, 2, 4 and 8 sec. One of the four methods was the usual method of reproducing an interval by two taps on a telegraph key. The three remaining methods, among which the method of stimulus-presentation was varied, required only one response for the reproduction of an interval. It was found that reproduction of a time-interval by only one response instead of two did not result in a significant change in either the accuracy or the consistency of time-judgments. For all four methods, the consistency of judgment decreased significantly from the 2- to the 4-sec interval and from the 4- to the 8-sec. interval. In contrast to the usual tendency to overestimate short intervals and to underestimate long intervals, the 8-sec. interval was significantly overestimated as compared with each of the shorter intervals. These results were compared with previous results, and it was concluded that time-judgments by the method of reproduction are greatly influenced by certain variations of experimental procedure.

PROFICIENCY IN FINGER-TRACKING AS A FUNCTION OF NUMBER OF FINGERS

By DONALD R. HOFFELD, SIDNEY SEIDENSTEIN, and W. J. BROGDEN, University of Wisconsin

One dimension of complexity in tracking a target with the fingers is represented by the number of fingers involved. The measurement of tracking proficiency for any given level of complexity thus defined requires the contribution of measures from all fingers, since the anatomical and other characteristics of the individual fingers suggest that there will be differences in tracking proficiency between the fingers. Thus the interaction of relative proficiency of the individual fingers and level of complexity is the general problem. Since proficiency is also a function of the amount of practice, any relations between complexity and proficiency might be expected to interact with the level of practice.

The present paper reports a series of four experiments designed to investigate the relation between tracking proficiency and complexity, and to determine the relative proficiency of the individual fingers in tasks of different complexity, over the course of practice sufficient to provide significant learning. The first two were concerned with tracking in a task of maximal complexity (four fingers simultaneously) and were designed to obtain learning curves as a function of the amount and distribution of practice. The remaining two experiments dealt respectively with one-finger tracking by each of the four fingers and with two-finger tracking of each combination of the four fingers. An experiment on three-finger tracking was planned but not completed because of the apparent lack of a consistent relation between the relative proficiency of each finger and the level of complexity.

METHOD

Apparatus. The apparatus provides for unidimensional compensatory tracking and consists of (1) a generator for the target-course, (2) four cathode ray-oscilloscopes (*CRO*), (3) four finger-tracking controls, (4) four clocks for scoring time-on-target, and (5) the necessary power supplies and circuits for coupling the components. The generator of the target-course consists of a metal cam driven by a 1-r.p.m. synchronous motor and four cam-followers located at 90° intervals. As the position

* Received for publication May 14, 1959. Supported in part by a grant from the National Science Foundation and by the Wisconsin Alumni Research Foundation.

of each follower is changed by the random contour of the cam, the shaft of a potentiometer is rotated. This rotation so changes the voltage applied to its associated *CRO* that the position of the target (a horizontal line) is moved up or down on the vertical axis. The target-course is the same for each of the four channels, except that the start and end are at different points in the total cycle. Fixed horizontal lines, one 5 mm. above and the other 5 mm. below the horizontal axis, define the on-target limits.

The tracking controls are molded rubber cups into which the first joints of the fingers are inserted. Each one of the four cups has a weighted cable attached to the tip. The cable runs over a pulley that rotates the shaft of a potentiometer. Movement of the appropriate finger up and down in a vertical plane thus provides a voltage change at the *CRO* to compensate the voltage from the target-course generator. The position of the target on the *CRO* is a display of error, representing the difference between the voltages from the target-course generator and the finger tracking-control. Thyratron circuits, operated by this voltage-difference, so control the clutches of the electric clocks that time-on-target is recorded in thousandths of a minute when the target is within the limits represented by the fixed horizontal lines on the *CRO*.

The arrangement of the display provides maximal compatibility of the stimulus-response. The four *CRO*s are controlled by the four fingers respectively from right to left, and movement of the control downward brings the target down. *S* sat in a comfortable chair that was so adjusted that the eyes of *S* were slightly above the horizontal axis of the display. The right arm rested on an adjustable platform with the palm of the hand down and the appropriate fingers were inserted in the tracking controls. Although only vertical movement of the control affected the position of the target, some horizontal movement of the fingers was unavoidable.

Before every experimental session for each *S*, the target was centered on the display, the scoring limits for the time-on-target area were checked, and the setting of the magnification of the error was calibrated. This latter adjustment provided 0.2 mm. of displacement of the target on the *CRO* for each millimeter of displacement of the tracking control; which is within the range for optimal proficiency in tracking.[1]

Subjects. The *S*s (180 in number) were student volunteers. All were right-handed.

Procedure. A trial consisted of one revolution of the course of the cam, *i.e.* 1 min. In Experiment I, 25 trials were made every session. In Experiments II, III, and IV, the number of trials was reduced to 10 per session. In all four of the experiments, measures of tracking proficiency were obtained for every *S* on 5 days with approximately 24-hr. intervals between successive sessions. In Experiments I and II, the task involved performance of the four fingers simultaneously. There were 15 *S*s in each of these experiments. In Experiment III, the task involved a single finger only and there were four groups of 15 *S*s each—one group for each of the four fingers. The task for Experiment IV involved the performance of two fingers simultaneously and there were six groups of 15 *S*s each for the six pairs of fingers.

The same instructions were given to every *S*. His task was to maintain the target

[1] W. F. Battig, E. H. Nagel, and W. J. Brogden, The effects of error-magnification and marker-size on bidimensional compensatory tracking, this JOURNAL, 68, 1955, 585-594.

within the two horizontal lines of each *CRO*. The clock for each finger recorded time-on-target only when the target was within the limits of the two horizontal lines. Information was also given *S* about the number of trials and experimental sessions to be held. Every *S* was given one preliminary trial to ascertain whether he understood the instructions. If he did not, further instructions were given.

Before every trial, *E* gave *S* a signal before pressing the switch to start the 1-min. course. The course-generator was shut off automatically. During the 30-sec. inter-trial interval, *E* recorded the time-on-target and reset the clocks to zero. There was a rest-interval of 2 min. between blocks of five trials. For the experimental sessions following the first, *S* was given information about his level of performance at the previous session and, when necessary, was reminded of instructions and procedure. Every attempt was made to have *S* work diligently and at his optimal performance.

RESULTS

The data for Experiment I revealed some decrement in performance toward the end of the 25-trial daily session. This was the reason for changing to 10 trials per session in Experiment II.

An orthogonol polynomial analysis of variance was performed on the data of Experiments I and II.[2] It revealed no significant differences between the two experiments, but there were significant differences between means for each finger, and significant linear and quadratic components in the over-all learning curves.[3] The mean times-on-target for each day in 0.001 min. per trial, averaged over all fingers, are presented in Fig. 1. The lines connecting the points are the best-fitting polynomial equations for each set of data. Comparable data and plots of the fitted equations are given also in Fig. 1 for Experiments III and IV.

As might be expected from an inspection of Fig. 1, an analysis of variance comparing the data for Experiments II, III, and IV demonstrates significant differences between the means of the three experiments. It also shows that the linear and quadratic components of the three learning curves differ significantly from one another. The Duncan Range Test shows that each mean is significantly different from each other mean.[4] The linear and quadratic terms of the polynomial equations for the learning curves (see Fig. 1) for one-finger and four-finger tracking are similar and are significantly different from chance. The terms for two-finger tracking are different from the one-finger and four-finger equations and the

---

[2] D. A. Grant, Analysis-of-variance tests in the analysis and comparison of curves, *Psychol. Bull.*, 53, 1956, 141-154. The orthogonal polynomial analysis of variance was used throughout this study.

[3] Since the assumption of homogeneity of variance was not satisfied by the data for this and all other analyses, the 1% level of significance is used in lieu of the 5% level for all tests.

[4] R. B. Duncan, A significance test for differences between ranked treatments in an analysis of variance, *Va. J. Sci.* 2, 1951, 171-189.

linear, quadratic, and cubic components are significant. Thus, the difference in distribution and amount of practice between Experiments I and II provides no significant differences in level of proficiency, amount acquired, or the form of the learning curve. Complexity is a significant factor, however, in that degree of proficiency increases from four-finger to two-finger to one-finger tracking. In addition, the form of the learning curves for each of these conditions of complexity is different in some degree from every other condition.

The analysis of variance of the data of Experiments II, III, and IV also showed significant differences between fingers and interaction of experiments and fingers. There was also a significant interaction of the slopes for each experiment and fingers. Separate analyses of variance were completed on the data of the three experiments, and equations were fitted to the learning curve for each of the four fingers for each experiment. Fig. 2 presents the empirical data and the equations grouped by fingers. The parts of the figure show the effect of complexity upon the acquisition of skill in tracking by each of the four fingers. In each case, proficiency increases as complexity decreases. The levels vary both as a function of the finger used and the complexity of the task. Moreover, the form of the learning curves for a given finger varies as a function of complexity. The form of the

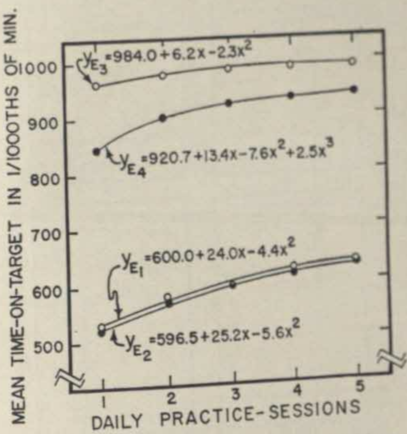


FIG. 1. THE RELATION BETWEEN COMPLEXITY OF THE TASK AND PROFIENCY

The notations for y (E_1, E_2, E_3, and E_4) refer to the data from Experiments I, II, III, and IV, respectively. Mean time-on-target in each case is for a single finger per trial. Four-finger tracking is represented by E_1 and E_2; two- and one-finger tracking by E_4 and E_3, respectively.

curve appears to be more closely related to level of complexity than to finger. This relation is demonstrated more clearly by Fig. 3 in which the curves of Fig. 2 have been regrouped by level of complexity.

The analyses of variance for the three experiments (II, III, and IV), show significant differences between means for each finger. The Duncan Range Test gives the following results: for Experiment II, the means for A(582.9) and D(533.8) are significantly different from all other means, but B(607.3) and C(617.6) do not differ significantly from each other; for Experiment IV, the means for B(884.3) and

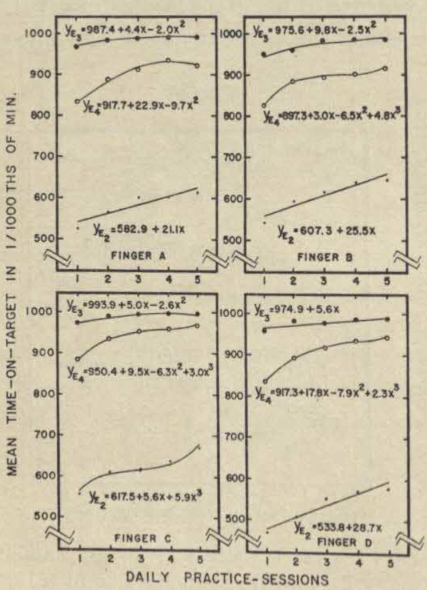


FIG. 2. THE RELATION BETWEEN COMPLEXITY AND PROFICIENCY BY THE DIFFERENT FINGERS

The notations for y (E_2, E_3, and E_4) refer to the data from Experiments II, III, and IV, respectively. Mean time-on-target is for a single finger per trial. Four fingers were used in E_2, one finger in E_3, and two fingers in E_4.

C(937.9) are significantly different from all other means, but A(898.3) and D(901.4) are not significantly different from each other; and for Experiment III, each mean is significantly different from every other mean (A = 983.5; B = 970.5; C = 988.1; D = 974.9). There is not a high degree of consistency in proficiency of fingers as a function of complexity. Finger C is the most proficient in Experiments III and IV, but in Experiment II its mean is not significantly different from that of B. Finger B shows the lowest degree of proficiency in Experiments II and IV, but in Experiment III this position is held by Finger D.

Although there are no significant differences in the form of the learning curves among fingers in Experiment II, the best fitting equations for Fingers A, B, and D

are linear, and that for Finger C has in addition a significant cubic component. For Experiment IV, the best fitting equations have significant linear, quadratic, and cubic components except for Finger A which has only significantly linear and quadratic components. The differences among fingers are not significant. The analysis for Experiment III shows significant differences among the fingers with significant linear and quadratic components for Fingers A, B, and C and only a significant linear component for Finger D.

The analysis of variance of the data for Experiment IV indicates that in addition to the results already presented there are significant differences between means

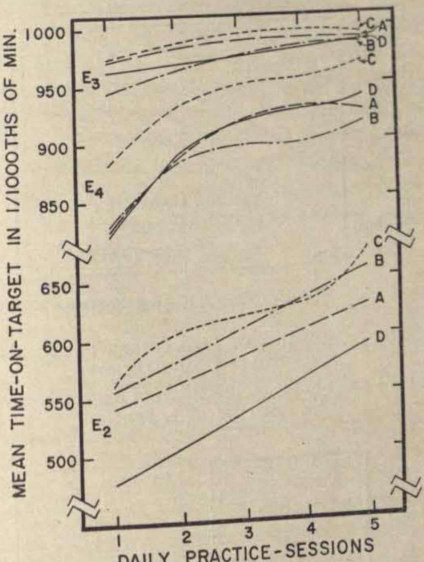


FIG. 3. THE RELATIVE PROFICIENCY OF FINGERS IN TASKS OF THREE LEVELS OF COMPLEXITY

for the combinations of two fingers, and for each finger as a function of the other finger involved in two finger-tracking. There are, however, no significant differences in form between combinations of two fingers or of single fingers as a function of the associated finger. The empirical data and fitted equations for the two finger combinations are presented in Fig. 4. The means for the combinations AD(863.7) and BD(873.6) do not differ significantly from each other but do differ significantly from all other means. The means for BC(906.6) and AB(907.7) do not differ significantly from each other but differ significantly from all other means except that AB does not differ significantly from that for AC(924.9). The mean for AC does not differ significantly from that for CD(942.9) which does differ significantly from all other means. Level of proficiency does not appear to be related to the spatial separation of the fingers (adjacent, separated by one, or by two) or to any other anatomical characteristics. Combinations involving Finger C appear to produce

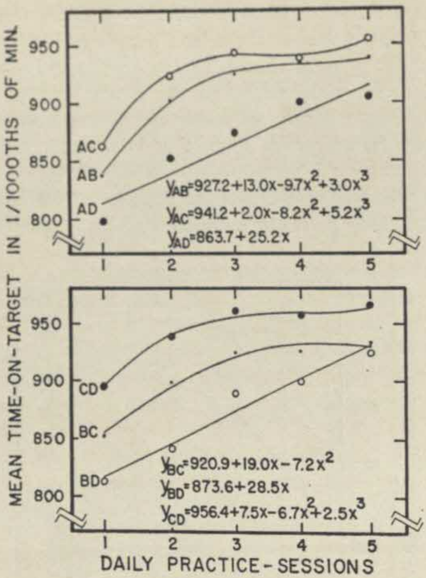


FIG. 4. THE ACQUISITION OF TWO-FINGER TRACKING IN THE SIX POSSIBLE COMBINATIONS

The notations for y refer to the combinations of fingers. Mean time-on-target is the mean per finger per trial.

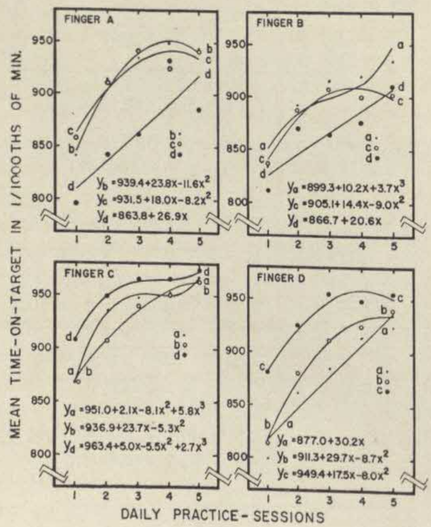


FIG. 5. THE ACQUISITION OF SINGLE-FINGER TRACKING AS A FUNCTION OF THE FINGER ASSOCIATED IN THE TWO-FINGER TASK

higher proficiency than any other combinations. Although there are no reliable differences in the form of the learning curves, combinations involving Finger D appear to produce linear curves.

The empirical data and fitted curves for single-finger tracking as a function of the finger associated in two-finger tracking are presented in Fig. 5. The mean for A*d*(863.8) is significantly different from A*c*(915.0) and A*b*(916.2) which do not differ significantly.[5] The mean for B*d*(866.7) is significantly different from B*c*(887.0) and B*a*(899.3) which do not differ significantly. The mean for C*d*(952.4) differs significantly from C*b*(926.3) and C*a*(934.9) which do not differ significantly from each other. The means D*a*(877.0) D*b*(873.8), and D*c*(933.4) all differ significantly from each other. The only consistent relation appears to be that when *d* is in combination with any other finger, the score for that finger will differ significantly in its score for any other combination. In two cases, A*d* and B*d*, there is minimal proficiency, and in one case, C*d*, there is maximal proficiency. Only in the case of Finger D does combination with other fingers provide significant differences and an apparently meaningful relation. As degree of separation between fingers increases, proficiency increases. There is some evidence for the existence of this relationship for the other fingers. Except for Finger C, minimal proficiency occurs at maximal separation of the fingers and in all cases, maximal proficiency occurs for adjacent fingers.

Discussion

Proficiency in tracking is inversely related to complexity of the task defined by number of fingers performing simultaneously. Although this relation is based upon experimental results only for one, two-, and four-finger tracking, it is highly probable that three finger tracking will be intermediate between the levels for two and four fingers.

The results clearly establish differential proficiency among the separate fingers in each of the combinations studied. These effects are relatively specific for each complexity level, except that the third finger (C) shows the maximal proficiency.

Relative proficiency in the two-finger task shows some consistencies when the performance of individual fingers is examined as a function of the associated finger. When the associated finger is the little finger (d), each of the fingers differs in its performance from that when it is associated with the other two fingers. In two cases, Fingers A and B, proficiency is minimal and for Finger C, proficiency is maximal. Since Finger D is presumably the weakest and least skilled finger, these results might suggest that *Ss* either concentrate on the other finger (A or B) or concentrate on Finger D to the detriment of Finger C. An alternative view is that proficiency of performance of Finger D varies directly with degree of separation between the associated fingers. Proficiency is maximal when the associate is c and decreases progressively with b and with a as associates. There is some evidence for this relationship among all fingers, since maximal proficiency is obtained from adjacent fingers and, with one exception (aC), minimal proficiency occurs with maximal separation of the fingers.

Summary

Four experiments are reported in which the proficiency of one-, two-,

[5] The designation 'A*b*' is used to indicate a score obtained from the forefinger (A) while it is associated in a task with the middle finger (*b*).

and four-finger tracking was measured in terms of time-on-target during the course of five daily practice sessions. Experiments I and II required four-finger tracking. They differed only in that the first required 25 trials per day, and the second 10. No significant differences were obtained either in the form of the learning curves or the level of proficiency achieved. Experiments III and IV required, respectively, one-finger tracking with each finger in turn, and two-finger tracking with each combination of two fingers.

Measures of proficiency decreased as complexity increased. The acquisition curves were curvalinear, but of different form in each experiment. Relative proficiency of the individual fingers varied as a function of the complexity of the task, but the only consistent relation over all levels of complexity was the superiority of the third finger (c). For two-finger tracking, the little finger (d) changed the proficiency of the associated finger to the greatest extent. For the little finger, the proficiency decreased as the separation of the associated finger increased. There was also evidence of this relationship for the other fingers in that maximal proficiency was obtained for adjacent fingers, and except for one case, minimal proficiency occurred with maximal separation of fingers.

CUTANEOUS DISCRIMINATION OF ELECTRICAL INTENSITY

By Glenn R. Hawkes, University of Virginia

The potentialities of the cutaneous senses for communication have been the subject of relatively little research as compared to the amount of work on such problems in vision and hearing. One of the aims of the present experiment was to provide data which would bear on the use of electrical cutaneous stimulation for purposes of communication. Most of the previous research which considered the possibility of cutaneous communication used mechanical, not electrical, stimulation of the skin.

Based on investigations of the discriminative limits for intensity, duration, and locus of mechanical vibration of the skin, a cutaneous system of communication has been devised by Geldard and his co-workers.[1] It was determined that three levels of intensity of stimulation, three durations of stimulation, and five loci were absolutely identifiable by their *O*s without error. Combinations of such cues were assigned meaning; training of *O*s demonstrated that these vibratory signals could be used for signaling purposes.

Although the system described above was concerned with sensations produced by the adequate stimulus, *i.e.* repetitive movement of the tissue by mechanical means, vibratory sensations can easily be elicited by the application of the inadequate stimulus, electrical current. An electrical cutaneous system of communication analogous to that using mechanical vibration is probably feasible, with the potential advantage that certain dimensions of electrical stimulation might prove to be more useful than their mechanical counterparts.

An investigation of the size of ΔI for electrical stimulation of the skin could perhaps provide information bearing not only on the above, but also on a more theoretical question: Would a determination of ΔI for stimulation by alternating current be a measure of the sensitivity of nerve or of receptor?

The punctate distribution of the skin's sensitivity may be explained by a suggestion made by Von Frey that particular kinds of cutaneous sensation depend on the action of particular kinds of specific neural terminations.[2] According to this view, pressure in hairy regions is subserved by the nerve-terminations at the follicles; pressure-sensations in hairless regions are said to be due to the action of Meisner's corpuscles or Merkel's disks.[3]

* Received for publication December 15, 1959. Based on a dissertation submitted to the University of Virginia in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer is indebted to Dr. J. F. Hahn and Dr. F. A. Geldard for their assistance during this research.

[1] F. A. Geldard, Adventures in tactile literacy, *Amer. Psychologist*, 12, 1957, 115-124.

[2] M. Blix, Experimentelle Beiträge zur Lösung der Frage über die specifische Energie der Hautnerven, *Z. Biol.*, 20, 1884, 141-156.

[3] F. K. Sanders, Special senses, cutaneous sensation, *Annu. Rev. Physiol.*, 9, 1947, 553-568.

Unfortunately, the neat explanation of the action of the peripheral nervous system by correspondence between sensitive spot and specific neural structure has not been confirmed. A few studies have reported high concentrations of encapsulated endings coincident with regions of high sensitivity.[4] The preponderance of evidence, however, has shown no such ending.[5]

The evidence does not preclude the possibility that end-organs, where they do occur, make a contribution to sensation. Further, there remains the possibility that nerves, or their terminals, may be maximally sensitive to different kinds of stimulation. A comparison of ΔI for adequate and inadequate stimulation may help to indicate whether or not end-organ or nerve is responding to alternating current applied to the skin. Such a comparison would be facilitated if data were available using similar methods of stimulus-presentation for the two kinds of stimuli.

Comparison with intensitive discriminations involving mechanical vibration of the skin is the primary interest of the present discussion, but it is desirable first to examine some results in audition, which have certain similarities to cutaneous vibration.[6] Auditory ΔI-values have been determined by two principal methods, the 'beat'-method, and the method of successive stimuli.

The 'beat'-method was introduced by Riesz, who avoided switching transients in a study of ΔI/I for pure tones by using two oscillators at slightly different frequencies to produce beats.[7] The differences in stimulus-intensity when beats just appeared or disappeared were used to measure ΔI, the smallest values being recorded with a beating rate of 3 per sec. (a rate of variation found optimal for the detection of beats in cutaneous mechanical vibration[8]). The stimulus-frequency, in the range of 200–1000 ~, was less important than the intensitive level of the standard in determining the size of ΔI/I, the Weber-fraction varying from about 1.24 to 0.15. He found that ΔI/I was larger at weak intensities of the standard than at stronger levels.

The method of successive stimuli was used by Dimmick and Olson[9] to determine the values of ΔI/I for pure tones. A short tone was presented to *O*, and, after a silent interval of 0.5 sec., a second tone to be compared in loudness with the first. Over the range of 256–1000 ~, ΔI/I was reported to vary from 2.63 to 0.91. The value of ΔI/I was reported to be larger at weaker levels of the standard when this method was used, as in the research reported above.

The necessity of avoiding transients led one investigator of ΔI/I for mechanical cutaneous vibration, Sherrick,[10] to the use of Riesz' beat-method. As in audition,

---

[4] B. Belonoschkin, Über die Kaltreceptoren der Haut, *Z. Biol.*, 93, 1933, 487-489; H. Strughold and M. Karbe, Vitale Färbung des Auges und experimentelle Untersuchung der gefärbten Nervenelemente, *Z. Biol.*, 83, 1925, 297-308.

[5] G. Weddell, Somesthesis and the chemical senses, *Annu. Rev. Psychol.*, 6, 1955, 119-136.

[6] G. von Békésy, Similarities between hearing and skin sensations, *Psychol. Rev.*, 66, 1959, 1-22.

[7] R. R. Riesz, Differential intensity sensitivity of the ear for pure tones, *Phys. Rev.*, 31, 1928, 867-875.

[8] J. A. Vernon, Cutaneous interaction resulting from simultaneous electrical and mechanical vibratory stimulation, *J. exp. Psychol.*, 45, 1953, 283-287.

[9] F. L. Dimmick and R. M. Olson, The intensive limen in audition, *J. acoust. Soc. Amer.*, 12, 1941, 517-525.

[10] C. E. Sherrick, Jr., Measurement of the differential sensitivity of the human skin to mechanical vibration, Unpublished M.A. thesis, University of Virginia, 1950.

the lowest values were recorded with a rate of variation of 2–3 per sec. Sherrick's results show no significant change due to frequency of stimulation, over a range of 250–1100 ~. His values for $\Delta I/I$, for vibration of the finger tip, were about 0.50 near the *RL*, and decreasing to 0.25 to 0.33 at higher intensities of stimulation. The relatively large values of $\Delta I/I$, compared to previous reports, are interpreted by Sherrick as reflecting more exact measures of sensitivity due to the absence of switching transients.[11]

The high values for $\Delta I/I$ obtained by Sherrick might have been due to attentional factors, *O*s' perhaps experiencing greater difficulty in attending to the stimulus than with more traditional methods of stimulus-presentation. If, however, this were the case, then Riesz' beat-method should have yielded values for auditory ΔI larger than those obtained by Dimmick and Olson with the method of successive stimuli; the reverse was true.

The method of successive stimuli was used to determine $\Delta I/I$ for mechanical cutaneous vibration by Schiller.[12] Two stimuli of 0.75-sec. duration, separated by a pause of 0.24 sec., were presented to the *O*s. The values of $\Delta I/I$ reported for a range of stimulus-frequencies from 100–800 ~ were between 0.11 and 0.14.

The determinations of $\Delta I/I$ which have been reviewed indicate that intensitive discrimination improves as the intensity of the standard becomes greater. With the use of different methods of stimulus-presentation, however, considerable differences in $\Delta I/I$ have been found both for audition and mechanical cutaneous vibration. Using the method of beats, Riesz found smaller values for auditory $\Delta I/I$ than Dimmick and Olson, who used the method of successive stimuli; however, Sherrick's values for vibratory $\Delta I/I$ with the beat-method are larger than those found by Schiller, who used the method of successive stimuli. An explanation of these differences might be better made if one used both methods within a single experiment, on the same group of *O*s.

The data reviewed above may be compared with the information available when the inadequate stimulus, an electrical current, is used. A study of the *DL* for intensity of tactual sensations aroused by alternating current at 23.6 ~ and 3035 ~ has been reported by Schöbel.[13] Naïve *O*s, stimulated on the hand, had values of $\Delta I/I$ of 0.02 to 0.03 at 23.6 ~, and 0.01 for 3035 ~, with a standard intensity of 115% of the current at *RL*. Reports of pain, fear of electrical stimulation, and other interfering sensations (such as those from muscle-contractions) prevented a determination of $\Delta I/I$ at higher intensitive levels of the standard.

Anderson and Munson[14] determined $\Delta I/I$ for alternating current applied to the forearm. At frequencies of 100, 500, and 5000 ~, $\Delta I/I$ was determined on two *O*s by the 'ABX' method; stimuli had a duration of 0.5 sec., separated by a 1-sec. pause. The average intensitive *DL* found with this procedure varied from 0.02 to 0.05 at a standard intensity of 562% of the *RL*-current, and did not change with

---

[11] V. O. Knudsen, "Hearing" with the sense of touch, *J. gen. Psychol.*, 1, 1928, 320-352.

[12] H. Schiller, Über die Amplitudenenunterschiedsschwellen des Vibrationssinnes beim Menschen, Unpublished Ph.D. dissertation, University of Erlangen, 1953.

[13] Edmund Schöbel, Versuche über Intensitätunterscheidung beim 'elektrischen' Tasten verschiedener Frequenzen, *Z. Sinnesphysiol.*, 66, 1936, 262-273.

[14] A. B. Anderson and W. A. Munson, Electrical excitation of nerves in the skin at audiofrequencies, *J. Acoust. Soc. Amer.*, 23, 1951, 155-159.

frequency. Intensitive discrimination was not reported for other intensive levels of the standard. These values were based on not one sensation, but on (a) a burning feeling near the surface of the skin, and (b) a deeper throbbing or pulsating pressure. Care was taken in the present work to determine values of $\Delta I/I$ for only a single kind of sensation, *i.e.* vibration.

The beat-method and the method of successive stimuli were used to determine $\Delta I/I$ in this experiment, for it would be desirable to know the limits of differential sensitivity independently of the method of stimulus-presentation. Frequencies of stimulation were selected which would facilitate comparison with available data for both the adequate and the inadequate stimulus; the frequencies selected were 100, 500, and 1500~. The reduction in size of $\Delta I/I$ as the intensity of the standard is increased seems to be a general characteristic of the results in audition and cutaneous vibration, but such data are not available for stimulation of the skin by alternating current. Standard intensitive levels were selected within the range eliciting sensations of vibration, but not pain.

Method

Apparatus. A block circuit-diagram of the apparatus is shown in Fig. 1. The switching amplifier[15] was designed to turn the signal on and off at rates of onset and offset which avoided the production of transients. The variable-inductance beat-former permitted the change of amplitude at a regular rate without affecting frequency or wave-form. The amplitude change resulted from changes in the impedance of an 11-Ω solenoid as its core was inserted or withdrawn. The core was attached to an eccentric cam fastened to the shaft of a variable-speed motor, which controlled the rate of movement of the core. A regulated power-supply with an output of 117-AC was used as the source of current for the equipment. The beat-former was used only for the determination of $\Delta I/I$ by the beat-method.

Procedure. Three *O*s were given preliminary training in the reporting of the sensations elicited by stimulation of the finger-tip by alternating current, and were also familiarized with the reporting of changes of intensity. The *O*s were able to discriminate fluctuations of intensity most accurately at a variation-rate of 1.5 to 2.5 per sec.; stimuli were varied in intensity at a rate of 2 per sec. when $\Delta I/I$ was determined by the beat method.

Current was delivered to *O* at a frequency of 100, 500, or 1500 ~, through an electrode 12 mm. in diameter resting on the pad of the index-finger. An electrode 25 mm. in diameter was in contact with *O*'s palm. Both electrodes were coated with Medcraft Electrode Paste.

Preliminary study of the size of $\Delta I/I$ for intensity of alternating current indicated that inter-session variability of *RL* for an individual *O* and differences of *RL*-values between *O*s were too great to permit use of the same absolute levels of current throughout the experiment. Accordingly, the measurement of ΔI required

[15] Built from a circuit-design kindly furnished to the Psychological Laboratory by Dr. W. A. Munson, of Bell Telephone Laboratories.

measurement of the *RL*, for each *O* on each day, and use of standard intensitive levels at values having a constant ratio to the *RL*. Intensitive levels selected were 120% and 200% of the current at *RL*.

When the beat-method was used, only the lower *DL* was determined; the beat-former provided only decrements from the standard intensities. The Method of Limits was used with a stimulus of 4.0-sec. duration.

The method of successive stimuli involved the same *O*s and standard intensitive levels as indicated above. Frequencies of stimulation were 100 and 1500 ~; 500 ~ was omitted since frequency made no difference in the results when the beat-method was used. A preliminary study indicated that no change in the size of ΔI/I

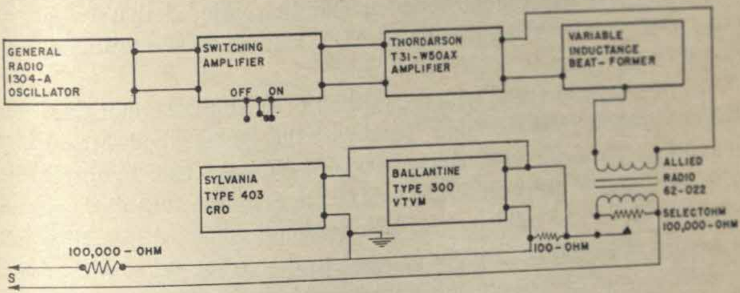


FIG. 1. BLOCK DIAGRAM OF APPARATUS

was to be expected from use of stimulus-frequencies of 100, 500, or 1000 ~; in addition, the upper *DL* was similar to the lower *DL*. The Method of Limits was used to determine the lower *DL* with two 2.0-sec. stimuli, separated by a 2.0-sec. silent interval. A counterbalanced order of presentation was used in both parts of the present experiment to reduce bias due to frequency of stimulation, ascending vs. descending series of the Method of Limits, and presentation of the standard first or the variable stimulus first with the method of successive stimuli.

RESULTS

Mean values for the three *O*s obtained by the beat-method are plotted in Fig. 2. The Sign-test indicated that the intensitive level of the standard had a significant effect on the size of ΔI beyond the 1%-level of confidence. Frequency of stimulation had a negligible effect on the size of ΔI. The mean values for the two intensitive levels of the standard for the three *O*s were 0.051 for ΔI/I at 120% of *RL*, and 0.035 for ΔI/I at 200% of *RL*. Individual differences tended to remain constant throughout the various standard intensity-levels and the several frequencies of stimulation.

There was no evidence for learning, *i.e.* the values showed no trend to be consistently lower for the later sessions than for earlier series of the same *O*s. No consistent difference was found between values recorded for ascending series and descending series within any one session.

Mean values of $\Delta I/I$ at various intensities and frequencies obtained by the method of successive stimuli are plotted in Fig. 3. The Sign-test indicated that the intensity of the standard had a significant effect on the size of $\Delta I/I$ beyond the 2%-level of confidence. It is noteworthy that one *O* (*RG*) exhibited a somewhat smaller $\Delta I/I$ with use of 1500~ at the weaker standard-intensity, and that another *O* (*BF*) appears to have a similar response at the stronger standard-intensity. Inasmuch as the data for any given *O* do not indicate such a trend consistently for the several frequencies and standard-intensities, it may be concluded that frequency had a negligible effect on the size of $\Delta I/I$. Preliminary observations tend to confirm this point.

The mean values found with the method of successive stimuli were not notably different from the values found with the 'beat'-method. Mean $\Delta I/I$ at 120% of *RL* was 0.053; $\Delta I/I$ at 200% of *RL* was 0.038. In agreement with the results of the 'beat'-method, $\Delta I/I$ values for the same *O* shown no consistent trend to be different for later sessions as opposed to earlier sessions, or for ascending series as opposed to descending series within any one session, or for series in which the variable stimulus was presented as the first member of the pair of stimuli vs. the series in which the standard stimulus was presented first.

Discussion

The most noticeable difference between the results of the present study and the results reported in audition and mechanical cutaneous vibration was the failure in the present instance to find a difference in the size of $\Delta I/I$ with different methods of stimulus-presentation. In the present study *O*s stated that the intensive changes with the 'beat'-method were difficult to perceive, and they reported a higher degree of subjective confidence when making judgments of intensitive changes with use of the method of successive stimuli. These observations might in part explain the failure to find differences due to method in this experiment but it does not resolve the conflict between the apparent superiority of the 'beat'-method in audition and its apparent inferiority in mechanical cutaneous vibration. It is not likely that attention would interact with modality in this manner.

The method of successive stimuli affords apportunity for a 'time-order' error, in which the second of two physically equal stimuli tends to be judged as greater than the first. It is unlikely that a serious time-order error appeared in the present study, for the variable stimulus was presented to *O* as the first member of the pair of stimuli in half of the series, and the standard stimulus was presented first in the remaining series. Time-

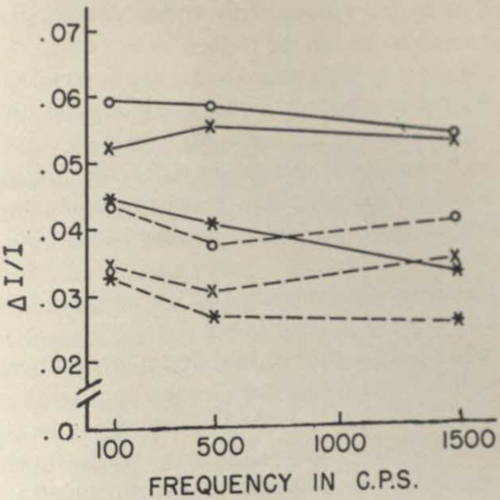


FIG. 2. MEAN WEBER-FRACTION FOR TWO STANDARD INTENSITIES, THREE FREQUENCIES, AND THREE *Os* BY THE BEAT-METHOD

---- Standard is 200% of the *RL*; —— Standard is 120% of the *RL*; o = Subject *BF*; x = Subject *GW*; * = Subject *RG*

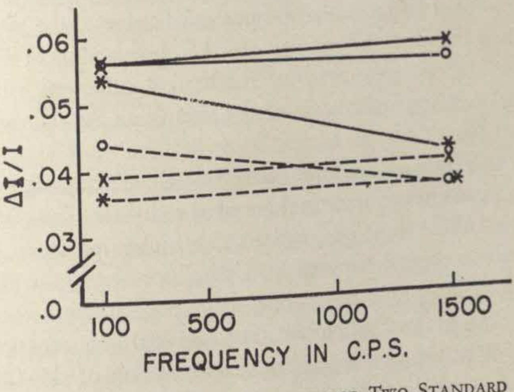


FIG. 3. MEAN WEBER-FRACTIONS FOR TWO STANDARD INTENSITIES, TWO FREQUENCIES, AND THREE *Os* BY THE METHOD OF SUCCESSIVE STIMULI

---- Standard is 200% of the *RL*; —— Standard is 120% of the *RL*; o = Subject *BF*; x = Subject *GW*; * = Subject *RG*.

order error might account for part of the difference between the values of $\Delta I/I$ reported in audition with the method of successive stimuli and the 'beat'-method. The large difference in the values reported, however, makes it unlikely that all of the difference could be so explained. A time-order error which may have appeared in Schiller's[16] data could have made his values of $\Delta I/I$ approach more closely those of Sherrick,[17] but Schiller's values are smaller, and therefore, the differences must be assumed to be due either to the use of different methods or to the possible presence of transients in Schiller's research.

A more probable explanation of the lack of differences between the results obtained by the two methods in the present study may be that the difference in values of $\Delta I/I$ reported by different investigators for mechanical cutaneous vibration, or for audition, is more apparent than real. These investigators used only a single method of stimulation in a given study. In the present study, the comparison of the two methods involved the same *Os*, equipment, and calibrations. Under these circumstances, the two methods did not produce different results.

The results of the present study are in agreement with studies of $\Delta I/I$ in other modalities in that the size of $\Delta I/I$ was larger at low than at high intensities of the standard.[18] The Weber-fraction was found to vary from about 0.052 at the weak intensity of the standard to about 0.036 at the strong intensity. These data do not agree with the predictions of Weber's law, that ΔI should be a constant proportional value of the standard. It is interesting to note that the reports of $\Delta I/I$ reviewed here indicate that $\Delta I/I$ is larger for audition than for mechanical cutaneous vibration, and that $\Delta I/I$ for electrical stimulation of the skin is smaller than either of the former.

The present values of $\Delta I/I$ are similar to those reported by Anderson and Munson even though these authors used a different method of presentation, *i.e.* the 'ABX' technique, and a much higher intensity of the standard.[19] Differences could have been anticipated in view of the fact that Anderson and Munson produced not only vibratory sensations by stimulating the forearm with alternating current, but also 'burning' sensations. Larger values of $\Delta I/I$ might be expected with stimulation of the forearm than with stimulation of the more densely innervated area used in the present study, *i.e.* the finger-tip. The use of a higher standard intensity may thus

---

[16] Schiller, *op. cit.* [17] Sherrick, *op. cit.*

[18] A. H. Holway and C. C. Pratt, The Weber-ratio for intensive discrimination, *Psychol. Rev.*, 43, 1936, 322-340.

[19] Anderson and Munson, *loc. cit.*

have served to make their $\Delta I/I$ values approximate the present data more closely.

The smaller values of $\Delta I/I$ found by Schöbel, compared to those of the present experiment, could be due to his use of naïve *Os*, some with sensory impairments, the production of more than one kind of sensation in his study, or differences in the psychophysical techniques used. Care was taken in the present work to avoid the production of switching transients, while Schöbel's apparatus may have produced such transients, which would have aided intensitive discrimination.

In the introduction to this paper, the question was posed of whether electrical cutaneous stimulation first acts on nerve or receptor. The small size of $\Delta I/I$ for electrical stimulation and the narrow sensory dynamic range, compared to those reported for adequate stimulation, suggest that the two kinds of stimuli act upon different neural mechanisms. This supposition receives support from Stevens' observations on the growth of loudness when current is applied directly to the auditory nerve. Stevens states that under these conditions, "The growth of loudness was many times steeper under electrical than under acoustical stimulation. The exponent of the power function under electrical stimulation . . . was of about the same order of magnitude as that observed when a 60-~ current was applied to the fingers."[20] Such a finding suggests that adequate stimulation of the skin normally acts first on receptors, while electrical stimulation acts directly on cutaneous nerves, as well as on receptors.

SUMMARY

The values of $\Delta I/I$ were measured for stimulation of the finger-tip by alternating current, at various combinations of stimulus-frequencies and standard intensities. Two methods of stimulus-presentation were used, both of which had previously been used to measure $\Delta I/I$ for audition and mechanical cutaneous vibration, the 'beat'-method and the method of successive stimuli. The values of $\Delta I/I$ differed negligibly for the two methods, in contrast to what had been found in audition and mechanical cutaneous vibration. Frequency of stimulation had no significant effect. Increases in standard intensity significantly reduced the size of $\Delta I/I$. The results indicate that stimulation of the skin by alternating current may act directly on cutaneous nerves, as well as on the cutaneous receptors.

[20] S. S. Stevens, Measurement and man, *Science*, 127, 1958, 383-389.

THE EFFECT OF SUBLIMINAL STIMULI ON GUESSING-ACCURACY

By WILLIAM SCHIFF, University of Miami

The evidence bearing on the behavioral effects of subliminal stimuli has been accepted by some psychologists and contested by others.[1] Whereas conflicting results have been obtained concerning the effects on several types of learning,[2] the evidence bearing on other types of behavior seems to point in an affirmative direction.[3]

Some recent negative results have cast doubt on the existence of subliminal phenomena.[4] The results of other investigations may be questioned on methodological grounds.[5] This study was designed, therefore, to test the phenomenon and to assess its generality under adequately controlled conditions.

The studies cited which yielded positive results suffer from several methodological weaknesses. These are: (a) the lack of separate independent control groups; (b) the use of small select samples of *O*s which make it difficult to generalize to a less sophisticated general population; and (c) inadequate determination of thresholds, and differing criteria of 'awareness.'

A control group is important in studies of subliminal perception for two reasons. First, a control group is more desirable than comparisons against mathematical levels of chance. When a control group is used, a deviation

---

* Received for publication September 30, 1959. The author is grateful to Dr. Benjamin H. Pubols, Jr. for his suggestions and criticisms, and to Dr. Robert B. Tallarico and Dr. Carl D. Williams for their aid in obtaining *O*s for the study.

[1] J. A. Vernon and D. H. Badger, Subliminal stimulation in human learning, this JOURNAL, 72, 1959, 265-266.

[2] P. A. Kolers, Subliminal stimulation in problem-solving, this JOURNAL, 70, 1957, 437-441; J. G. Miller, The role of motivation in learning without awareness, this JOURNAL, 53, 1940, 229-239.

[3] Miller, Discrimination without awareness, this JOURNAL, 52, 1939, 562-578; L. E. Baker, The influence of subliminal stimuli upon verbal behavior, *J. exp. Psychol.*, 20, 1937, 84-100; R. M. Collier, An experimental study of the effects of subliminal stimuli, *Psychol. Monogr.*, 52, 1940 (No. 236), 1-59, A. C. Williams, Jr., Perception of subliminal stimuli, *J. Psychol.*, 6, 1938, 187-199.

[4] Vernon and Badger, *op. cit.*, 266.

[5] W. F. Day, Serial non-randomness in auditory differential thresholds as a function of interstimulus interval, this JOURNAL, 69, 1956, 387-394; W. S. Verplanck, G. H. Collier, and J. W. Cotton, Non-independence of successive responses in measurement of the visual threshold, *J. exp. Psychol.*, 44, 1952, 273-282; L. D. Goodfellow, The human element in probability, *J. gen. Psychol.*, 23, 1940, 201-205.

from chance may not be attributed to procedural or subjective variables; thus the probability of an experimental error is reduced. On the other hand, if such variables affect control *O*s in a direction *opposite* to that for the experimental *O*s, an *empirical* control for chance would reduce the probability of a Type II error. Secondly, a separate control group takes account of such perceptual or response sets as might develop in an initial testing session and be carried into a second session. The need for a separate control group is particularly evident in view of recent evidence for non-random distribution of serial responses.[6]

Whereas the studies of Miller and Williams[7] used a control series (although no separate group was used), other investigations have depended entirely on classical chance for purposes of comparison.[8] The more recent studies which have used separate control groups, were concerned with the more complex process of problem solving, used different procedures, and obtained conflicting results.[9]

All of the older studies cited have used very small samples (usually four to ten) which consisted in most cases of graduate students in psychology. In one study, *O*s were *selected* on the basis of high scores in initial sessions.[10] Such samples may be adequate in the early stages of investigation but should now be expanded.

The method used in the determination of the threshold is critical because it entails a criterion of awareness as well as the connotations of the statement that the stimulus is below 'awareness.' In the studies of Newhall and Sears, and Kolers, the usual 50%-limen or no measurement at all was used. Vernon and Badger used a 25%-criterion, and Collier used an 'accuracy criterion.'[11] The same criterion has been used in the more recent studies of 'subception.' In these cases subliminal meant 'sometimes aware.' In the investigations of Baker, Miller, and Williams, however, the 'absolute onset' criterion, *i.e.* the lowest measured threshold, was so used that subliminal meant 'completely unaware' of the stimulus-figure, or inaccessible to verbal report. None of these criteria coincides in phenomenal experience. If criteria of 50% or 25% are used, *O* can sometimes 'see' parts of the figure and thereby increase the accuracy of his guesses through

[6] Day, *op. cit.*, 394; Verplanck, Collier, and Cotton, *op. cit.*, 280; Goodfellow, *op. cit.*, 203.

[7] Miller, *op. cit.*, 1939, 568; Williams, *op. cit.*, 191.

[8] Baker, *op. cit.*, 95.

[9] Kolers, *op. cit.*, 441; Vernon and Badger, *op. cit.*, 265-266.

[10] Williams, *op. cit.*, 190.

[11] I. Goldiamond, Indicators of perception: I. Subliminal perception, unconscious perception—an analysis in terms of psychophysical indicator methodology, *Psychol. Bull.*, 55, 1958, 373-411.

information of which he is aware. The use of the 'accuracy criterion' enables an individual to respond verbally at chance levels while 'seeing' the stimulus but remaining unable to distinguish its differentiating features. On the other hand, he may receive sensory information from such stimuli, and respond in excess of chance.[12] This criterion then, corresponds to the accuracy of the discriminative response and is not necessarily related to phenomenal awareness, whereas the usual psychophysical criteria define 'subliminal' in a statistical, but not a phenomenal sense. To resolve some of these difficulties, the following experiment was performed.[13]

METHOD

Observers. Sixty-nine *O*s (41 men and 28 women) volunteered for the experiment. Most of the *O*s were students of introductory psychology with a maximum of two courses in psychology. All were students, but since they were enrolled in summer courses, the sample was more heterogeneous than the usual population of students. Many *O*s were teachers and adults from other professions and occupations, the age-range being 17–53 yr. with a mean of 24 yr. The *O*s enrolled in courses received class credit for participation.

Apparatus. Five geometric figures (a circle, square, triangle, diamond, and a square having a cross within it) were drawn with India ink on white 11 × 14-in. cards. The height and width of the figures were 2 in., and the line-thickness 0.05 in. These drawings were then copied on transparent-base film to produce positive 35-mm. transparencies on which the figures measured 3/16 in. in height and width. A blank slide was prepared for the control series.

A 35-mm. projector (S.V.E. Model SM-1) with a standard 300-w. lamp was fitted with an S.V.E. *Speed-i-o-scope* shutter. The lamp-voltage was controlled by a Variac. The leading edge of the lens of the projector was 17 ft. from a Radiant beaded-glass wall-screen. The *O*'s chair was placed 13 ft. 9 in. at center from the screen, and 15 in. to the right of the center of the line of projection. The projected image measured 8 7/16 in. in height and width, and ⅛ in. in line-thickness. The center of the image was 5 ft. from the floor.

The projection-room measured 31 × 18 ft. and was illuminated by a single shielded 100-w. bulb situated in a ceiling fixture 6 ft. behind the projector. The brightness of the screen as measured with a Macbeth Illuminometer 12 in. from the screen is shown in Table I.

Procedure. The *Os* were randomly assigned to experimental and control conditions by *E*'s flipping a coin before *O* entered the room. They were individually tested at night to insure fairly constant illumination. The *O*s were seated and given the following instructions.

---

[12] P. O. Bricker and A. Chapanis, Do incorrectly perceived tachistoscopic stimuli convey some information?, *Psychol. Rev.*, 60, 1953, 181-188.

[13] The criterion of 'absolute onset' of awareness, as used by Baker, Miller, and Williams, was the criterion used by the author, with additional verbal checks and a safety margin of 2 v. 'Subliminal' was, therefore, operationally defined as a point below the lowest of 16 threshold measures, in which the *O* reported when he was first aware of any part of the stimulus-figure.

We are trying to determine whether the accuracy of guesses can be influenced by showing the guesser what it is he is guessing—at levels of illumination so low he is not aware of the stimulus. In this case the stimuli will be some projected geometric figures, and you will be guessing which one is being shown. First, I would like to determine at what point you become aware of these figures. I will show them to you first at high illumination, then reduce the illumination so that you cannot see the figure. Then I will gradually increase the illumination—each time snapping the shutter twice. As soon as you first become aware of the figure or any portion of it, please say 'OK.' We shall repeat the procedure several times.

The Variac was first set at 60 v. and the figure shown twice in rapid succession at 0.01 sec. The figure was named (*e.g.* "this is the circle"), the Variac was reset at 20 v., and 1-v. increments made at 3-sec. intervals (per pair) up to threshold. Three such measurements were made for each figure. The *O*s were then instructed as follows:

Now I am going to show you a random series of these figures. By this I mean they are in no logical pattern, the figures may or may not appear an equal number of times, and they may or may not appear several times in succession. I would like you to look at the screen just as you did before, and each time I snap the shutter twice, *guess* which of the five figures is being shown. Should you be aware of anything on the screen other than a dim flash of light, please tell me right away. Only if the guess is a *pure* guess is it valid. Simply guess the first figure that comes to mind after the shutter has been snapped. Remember, the five figures are the circle, square, diamond, triangle, and window.

TABLE I

ILLUMINATION AND CORRESPONDING SCREEN BRIGHTNESS

Illumination (v.)	0	20	25	30	35	40	60
Brightness (ft.-C.)	1.5	1.6	1.9	2.4	3.2	3.6	15.0

The Variac was set 2 v. below the lowest measured threshold for each *O*, and a verbal warning was given each time before the pair of exposures was made. The interval between each pair of 0.01-sec. exposures was about 10 sec. The experimental group, Group E, was shown the series T T C D W S D S S T D C C S W T S W W T C D D C W and their guesses were recorded. The control group, Group C, was shown a blank slide which was removed from the projector and replaced in the same manner as the slides for Group E. The guesses of Group C were recorded against the same series noted above, yielding the empirical chance-scores. Upon completing the series of 25 guesses, the *O*s of Group C were shown five slides with figures, on which they guessed, and were then asked if they noticed anything different about the last five slides. Upon completing the last guess, all *O*s were asked if they were ever aware of any of the figures or parts thereof. They were then given their results and an accompanying explanation of the experiment.

A one-tailed test predicting superior accuracy for Group E was decided upon prior to the experiment, and was used to test the null hypothesis. The point of rejection was set at 0.01, and the 1% level was preselected for other tests of significance.

RESULTS

Awareness. Two *O*s in Group E reported cues received from negative after-images, and their data were excluded from the results. None of the

other 69 *O*s reported any awareness of the figures, even after extensive questioning. Members of Group C reported no difference for the terminal comparison-series, except that a few reported that the last five slides seemed brighter. Many *O*s in both groups commented that they could not see anything during the guessing-sequence, and were instructed by *E* to continue guessing.

Distribution of guesses. To provide comparable results for the unequal groups (37 *O*s in Group E and 32 *O*s in Group C), the frequencies of guesses were converted to proportions, which are shown in Table II. These figures indicate an approximately even distribution of guesses, although the square was guessed somewhat more frequently by both groups. When

TABLE II

ANALYSIS OF DISTRIBUTION OF TOTAL GUESSES AND CORRECT GUESSES

Figure	Proportion of all Guesses		Proportion correct Each Figure	
	Group E	Group C	Group E	Group C
C	.20	.20	.30	.18
S	.26	.25	.28	.18
T	.19	.18	.21	.15
D	.17	.17	.36	.19
W	.18	.20	.35	.19

the results for all the figures were combined, the number of experimental *O*s scoring above classical chance (five correct) as compared to the corresponding number of control *O*s, was significant in the predicted direction, with $\chi^2 = 15.00$ ($d.f. = 1, P < 0.001$).[14]

Mean differences. When the number of correct guesses for each *O* was treated as a raw score, an approximately normal distribution resulted. The mean for Group E was 7.49, with an *SD* of 2.51. The mean for Group C was 4.69, with an *SD* of 2.05. The resulting *CR* was 4.99 ($P < 0.00001$). Replacing the obtained mean of Group C with the mean expected according to classical chance did not appreciably alter the results.

Runs. The *Os* of Group C tended to guess quite often in *runs* of two, three, or more identical guesses. The choice of figure was not consistent between groups as shown in Table II. To test the significance of this observation, which might bear on the explanation of the subliminal effect, a run-score was computed for each *O*. This score was taken as the sum of all guesses which were part of two or more repeated guesses. The means

[14] The size of samples was conservatively equated by adding the difference to the 'above-chance' category of the control group.

were 6.73 for Group C, and 3.77 for Group E. The obtained *CR* of 3.77 permits rejection of the null hypothesis at the 1/10% level.

Discussion

Accuracy of guessing. The results of this investigation give strong support to the hypothesis that visual stimuli may influence the verbal behavior of *O*s who are unaware of any sensory images or information. The modest but rather consistent superiority of Group E suggests that such subliminal sensory information can be used by *O*s.

The results support the findings of Baker, Collier, Kolers, Miller and Williams,[15] under conditions which control for response-bias and other possible vitiating factors by the use of a separate control group. The generality of the phenomenon was extended by the use of a larger and more heterogeneous sample. The results, however, are contrary to those reported by Vernon and Badger.[16] This discrepancy may have resulted from differences in the mode of presentation of the stimuli, or differences in the complexity of the responses. It is more likely, however, that the lack of agreement is due to the factor of attention, since Vernon and Badger's *O*s apparently were not attending to the exact location of the projected stimuli. Furthermore, those *O*s apparently did not even know such stimuli were being projected, and hence were not 'looking for' the stimuli. Although Miller's *O*s did not know there were physical images being projected on the mirror at which they were staring, they were looking at the point of projection supposedly for purposes of *ESP*. In most of the other studies cited, the *O*s knew they were looking at subliminal images.

Patterns of guesses. The patterns of guesses indicated that in the absence of subliminal sensory information *O*s tended to guess in runs. Such perseverative behavior apparently was determined by idiosyncratic factors, although the square may have been favored by both groups because of its similarity to the window, its familiarity, or the fact that the flash of light was rectangular, conforming to the shape of the aperture of the slides. Day's interpretation of the serial effect was not supported, because there was no consistency between groups in the choice of figures in the runs, even though the same procedure was followed with both groups.[17] Day's interpretation apparently does not apply to all types of serial behavior.

[15] Baker, *op. cit.*, 99; Collier, *op. cit.*, 34; Kolers, *op. cit.*, 441; Miller, *op. cit.*, 1939, 598; Williams, *op. cit.*, 199.

[16] Vernon and Badger, *op. cit.*, 266.

[17] Day pointed to the possibility that serial effects were attributable to unique conditions of stimuli in the serial procedure, *op. cit.*, 388.

Goodfellow's analysis of guessing responses—indicating a preference for alternation, asymetry and avoidance of runs—was supported only by the asymmetry of the responses of Group C.[18] Group E's patterns, however, were in accordance with Goodfellow's findings.

While perseveration of responses may account for the failure of Group C to reach classical chance, it cannot account for the highly significant accuracy of Group E above classical chance. The non-independence of the responses of Group C confirms the need for an empirical control group in studies of serial verbal behavior.

SUMMARY

This study was designed to test the effect of subliminal visual stimuli on verbal behavior under controlled conditions and with a heterogeneous group of *O*s.

Five geometric figures were tachistoscopically presented to an experimental group at a speed and illumination below their lowest measured absolute threshold. Blank slides were presented to a control group under the same conditions. Analysis of the data revealed that the accuracy of the experimental group significantly exceeded that of the control group. In addition, there was a significantly greater tendency for the control group to guess in runs.

---

[18] Goodfellow, *op. cit.*, 202.

TACTUAL VERNIER-ACUITY

By P. J. Foley and E. V. T. Dewis, Defence Research Medical Laboratories, Toronto, Canada

Knowledge of tactual acuity is limited chiefly to the two-point threshold, the minimal perceptible separation between two adjacent stimuli. This threshold varies with the part of the body tested, but the tip of the index-finger, the most sensitive, shows minima of from 1 to 2 mm. Tactual vernier-acuity, the minimal perceptible displacement of one edge with respect to another continuous edge, had not been investigated. The present study was undertaken to supply data on this aspect of tactual sensitivity.

The diameter of the rods used in determining thresholds and the separation between these rods, were considered to be possible sources of variance, and were treated as experimental variables.

Method: *(1) Apparatus.* The main component of the apparatus is a micrometer stage for a travelling microscope giving direct readings of lateral displacement to 0.01 mm., and vernier-readings to 0.001 mm. The amount of lateral displacement is controlled by the large drum, one complete revolution of which causes the platform to move 0.5 mm. The upper rod or target is fixed in position, and the lower rod, mounted upon the stage, moves relative to it.

The targets, cylindrical in shape, were pieces of standard drill-rods of diameters 0.125 in., 0.250 in., and 0.500 in., respectively. A 50-mm. length, selected for parallelism, was cut from each rod. These lengths were then cut horizontally through the center, and the diameters of each half at the point of contact measured. The measurements for the rods finally selected were 3.162 mm., 6.337 mm., and 12.675 mm., for each half, respectively.

Four steel-gauge blocks were machined to thicknesses of 1 mm., 2 mm., 4 mm., and 7 mm. Accurate vertical separations between target-rods could then be made by placing the appropriate block or combination of blocks upon the lower rod and allowing the upper rod to rest upon the gauge.

Determination of the point of physical equality, the point at which the axes of the two target-rods were aligned precisely, was achieved by projecting a shadow-graph of the rods, enlarged eight times, upon a ruled screen, and making a visual match. This procedure was used for each diameter. In all cases the point of physical equality was found to be at a scale setting of 10.300 mm.

Procedure. Three conditions of rod-diameter *i.e.,* 3.162 mm., 6.337 mm., and 12.675 mm., were combined with five separations between rods, *i.e.* 0, 1, 3, 5, and

* Received for publication November 10, 1959. From the Defence Research Medical Laboratories, Project No. 212, DRML Report No. 212-10, PCC No. D77-38-50-04, H.R. No. 186.

7 mm., to give a 3×5 factorial design. The resulting 15 combinations were presented to each *S* in three sessions. During each session the effect of the five separations for only one rod-diameter was investigated. Twelve consecutive determinations were made for each of the five separations. The first 2 determinations in each case were treated as practice, and the analysis was based on the last 10. The order in which the different separations were presented within a session was random.

Six *S*s (all women), who were paid for their services, were used. Each was allocated one of the six possible orders of presentation of the three conditions of rod-diameter.

For determination of a single threshold, *S* was seated before the apparatus which was clamped to a table. A screen was so placed between *S* and the apparatus that no visual cues could be used. The lower rod was then displaced by *E* with respect to the upper by at least 1 mm. Displacement was alternately to the right and left. The tip of *S*'s left index-finger was placed against the rods, only vertical movement of the finger being permitted. *S* grasped the drum with the other hand and adjusted the relative position of the rods until no displacement could be perceived. No time limit was imposed.

Results. In the treatment of the results a distinction is made between the accuracy of *S*'s settings, and the precision of these settings. Accuracy refers to the degree to which the point of subjective equality approaches the point of physical equality. The index of accuracy is the difference between the mean setting and the point of physical equality. Precision refers to the degree to which each of 10 consecutive settings for a given condition approaches the point of subjective equality for that condition. The index of precision is the standard deviation. In general, investigators of visual vernier-acuity use only an index of precision as the acuity-threshold.[1] The logic here is that high precision implies a capacity to detect very fine differences. This view is meaningful only if one assumes that high precision and low accuracy cannot co-exist. The micrometer-stage has a range of displacement of from 0 to 25 mm. *S*'s settings were measured in these scale units, *e.g.* a setting of 10.000 would indicate that the lower rod was displaced 0.300 mm. to the left of the upper rod; a setting of 10.600 would indicate a displacement of 0.300 mm. to the right. In the analysis of precision the values used are the actual scale-settings. In the analysis of accuracy, the values used are the differences between the scale-settings and 10.300. A minus-value indicates that the lower rod is to the left, a plus-value that it is to the right.

Table I details the results of the experiment. In order to assess the effects on the precision of tactual vernier-acuity of variations in rod-diam-

[1] R. N. Berry, Quantitative relations among vernier, real depth, and stereoscopic depth acuities, *J. exp. Psychol.*, 38, 1948, 708-721; The relation of vernier and depth-discrimination to width of test rod, *ibid.*, 40, 1950, 520-522.

TABLE I

ACCURACY AND PRECISION OF TACTUAL VERNIER-ACUITY
(Mean displacement in mm. from point of physical equality and *SD* about point of subjective equality.)

Diameter of rod (mm.)	Separation between rods: 0		1		3		5		7	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3.162	−0.008	0.024	−0.025	0.039	0.111	0.113	0.367	0.239	0.669	0.316
6.337	−0.027	0.032	−0.023	0.044	0.072	0.083	0.195	0.218	0.282	0.400
12.675	−0.059	0.038	−0.034	0.058	−0.016	0.109	0.072	0.179	0.395	0.344

eter and separation between rods, an analysis of variance of the log variances was carried out. Each value was first coded by multiplying by a factor of 10,000 to eliminate negative characteristics. Varying the diameter of the rods does not affect the precision of tactual vernier-acuity. Varying the separation between the rods has however, a marked effect ($F = 57.194$, $df. = 4$ and 20, $P < 0.001$). *S*-differences are, of course, significant ($F = 23.826$, $df. = 5$ and 40, $P < 0.001$). The interaction of *S*s with separations is also significant ($F = 3.269$, $df. = 20$ and 40, $P < 0.01$). To investigate this effect further, the results for each *S*, rod-diameters pooled, were plotted and are shown in Fig. 1. It can clearly be seen that for all *S*s pre-

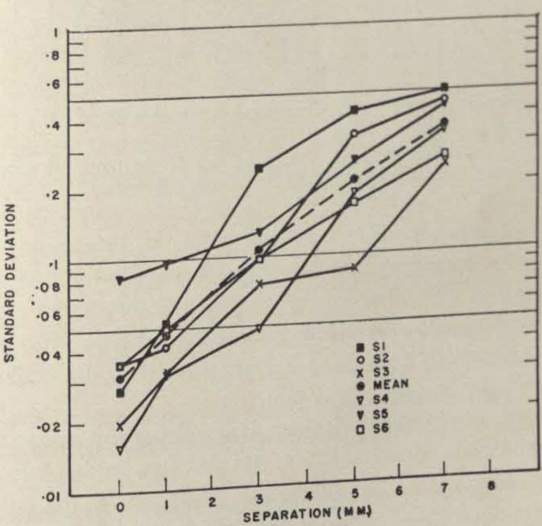


FIG. 1. PRECISION OF TACTUAL VERNIER-ACUITY AS A FUNCTION OF THE SEPARATION BETWEEN THE TARGET-RODS

(Precision is expressed as the logarithm of the standard deviation of the settings about the point of subjective equality.)

cision decreases as the separation between target rods increases. The interaction is due to the fact that the *S*s vary in the rate at which the decrease takes place. This slight variation does not affect the main conclusion that the precision of tactual vernier-acuity, expressed as the logarithm of the standard deviation, decreases linearly with increases in separation between the target-rods at least to a value of 3 mm. Analysis of the accuracy of tactual vernier-acuity is complicated by the finding that variance increases with separation. Less reliance can be placed on the means for the larger separations. To facilitate comparisons, the mean error for each *S* for each

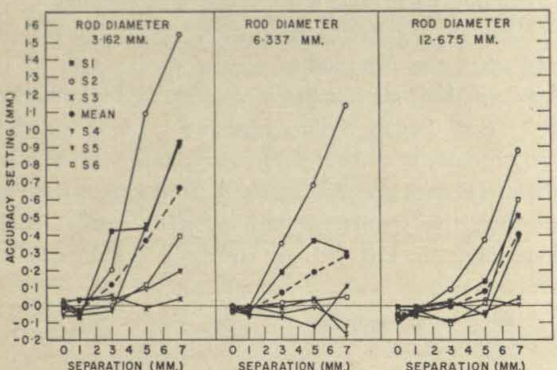


FIG. 2. ACCURACY OF TACTUAL VERNIER-ACUITY AS A FUNCTION OF THE SEPARATION BETWEEN RODS

(Accuracy is expressed as the deviation (mm.) of the point of subjective equality from the point of physical equality.)

rod-diameter at each separation is shown in Fig. 2. In general, beyond a value of 1 mm., increasing separation between target-rods decreases the accuracy of setting. This effect, however, is much more marked for some *S*s than for others. Thus, S_3 shows almost no effect, whereas S_2 shows a very large effect. Generally, however, no difference can be said to exist between no separation and a separation of 1 mm.

The effect of increasing rod-diameter is also dependent upon separation. Up to a separation of 1 mm., diameter has no effect, but beyond that value there seems to be a trend toward increased accuracy for the larger diameters as separation increases. The mean curves for the two larger diameters show no significant difference from each other. Pooling the results for no separation for the three rod-diameters gives a mean error of − 0.031 mm., and for 1 mm. separation, − 0.027 mm.

Discussion. It is interesting to note that just as visual vernier-acuity is more sensitive than visual two-point acuity, so also is tactual vernier-acuity more sensitive than tactual two-point acuity. The fact that rod-diameter has no effect on precision is analogous to Berry's finding that rod-width has no effect on visual vernier-acuity.[2] The effect on accuracy, however, is surprising. At small separations, 0 and 1 mm., rod-diameter has no effect. Beyond these values the two larger diameters show much less effect of separation than the small rod-diameter. This finding implies that summation is a factor in accuracy of setting, but that there is a threshold-value of separation below which summation has no effect, and an upper limit beyond which increments in rod-diameter do not increase the summation-effect. In other words, a gap up to 1 mm. is easy to bridge; beyond that point some lateral assistance is required. This assistance, analogous to propelling an orange pip by squeezing it, is only useful up to a certain point. It is unfortunate that Berry did not report his *Ss*' accuracy of setting. The differential effect of rod-diameter on precision and accuracy may also obtain in the visual modality.

There seems to be a significant trend for the *Ss*, at the two separations where precision and accuracy are greatest, to displace the lower rod to the left of the upper one. This can be seen in Fig. 2. This consistency of behavior suggests that it is easier to detect discontinuity by moving the finger down into the displaced edge, although, of course, the reverse might be the case.

It was thought that the generally consistent differences between *Ss* might be due to differences in finger-tip dimensions, a small tip being less accurate at large separations. Measurement of the width of the finger-tip held against a rod under light pressure, was made for four *Ss*, the other two being no longer available. The dimensions were 7.40 mm. for S_4, 9.75 mm. for S_2, 10.30 mm. for S_1, and 10.90 mm. for S_3. Reference to Fig. 2 will show that the differences among the *Ss* in accuracy cannot be attributed to differences in this dimension.

Summary. An experiment was carried out to determine the threshold for tactual vernier-acuity as affected by the diameter of the test-rods, and the separation between them.

The precision of setting is not affected by the diameters of the test-rods, but by the separation between the rods. Precision decreases as separation increases. Under optimal conditions, *i.e.* at zero separation, the value for

[2] Berry, *op. cit.* 1950, 520-522.

precision, expressed as the standard deviation about the point of subjective equality, is ± 0.031 mm.

The accuracy of setting is affected by the diameters of the test-rods and their separations, which interact with each other. The same accuracy is found at separations of 0 and 1.0 mm. for all diameters. Beyond these values accuracy decreases more rapidly for the smallest diameter, as separation increases, than for the two larger diameters, between which no difference in accuracy is noted at any separation. The mean value for accuracy, expressed as the difference between the point of physical equality and the point of subjective equality, is -0.031 mm. This value is the average of the values for the three rod-diameters at zero separation. The standard deviation about this mean is, of course, the associated index of precision, ± 0.031.

THE EFFECT OF AUDITORY STIMULATION UPON THE CRITICAL FLICKER FREQUENCY FOR DIFFERENT REGIONS OF THE VISIBLE SPECTRUM

By BARBARA MAIER and WILLIAM BEVAN, Kansas State University, and ISAAC BEHAR, VA Hospital, Durham, North Carolina

Despite the appearance from time to time during the past seventy-five years of reports on sensory interaction, these relationships have been little more than an obscure problem of tangential interest to Western psychologists.[1] The term 'inter-sensory effect' has been applied to a variety of phenomena: some related to associative memory, others reflecting the operation of attitudinal variables or set, still others that are classed as purely sensory events.[2]

The ambiguous status of this area is probably attributable to the prevalence of conflicting empirical reports[3] and to the absence of an adequate conceptual framework within which to accommodate positive findings. Sensory psychology has been marked by a peripheralism, probably the result of a misinterpretation of the doctrine of specific energy of nerves,[4] that, short of an assumption of centrifugal excitation, renders any consideration of intermodal effects meaningless; but there have been changes in the idiom of both psychology and physiology—*e.g.* Lorente's principle of reciprocity, Hebb's concepts of cell assembly and phase-sequence, Werner and Wapner's sensory-tonic theory of perception, Helson's adaptation-level, and Magoun's studies of the reticular activating system—that make a meaningful approach to the phenomenon more accessible. Still, these can, for the moment, only provide a congenial climate for investigation. None allows confident prediction of the wide variety of specific effects that have been reported. Explanations, where they are found—Kravkov's hypothesis concerning the effect of sound on visual acuity constitutes an example[5]—are formulated, for the most part, in terms of the particulars of the test-situation.

* Received for publication November 13, 1959. This study was conducted at Emory University.

[1] Not so with Russian investigators. London, in his review of the Soviet literature (I. D. London, Research on sensory interaction in the Soviet Union, *Psychol. Bull.,* 51, 1954, 531-568), includes over 500 items, most dated after the middle 1930s.

[2] H. F. Gaydos, Intersensory transfer in the discrimination of form, this JOURNAL, 69, 1956, 107-110; I. L. Child and G. R. Wendt, The temporal course of the influence of visual stimulation upon the auditory threshold, *J. exp. Psychol.,* 23, 1938, 109-127; S. V. Kravkov, Ueber die Beeinflussung der Unterschiedsempfindlichkeit des Auges durch Nebenreize, *Arch. Ophthal.,* 12, 1932, 105-11.

[3] Alphonse Chapanis, R. O. Rouse, and Stanley Schachter, The effect of inter-sensory stimulation on dark adaptation and night vision, *J. exp. Psychol.,* 39, 1949, 425-437.

[4] F. Hayek, *The Sensory Order,* 1952, 10 ff.

[5] S. V. Kravkov, Changes of visual acuity in one eye under the influence of the illumination of the other or of acoustic stimuli, *J. exp. Psychol.,* 17, 1934, 805-812.

Among the most intriguing, and also the most disturbing, of the reported intermodal relationships are the 'pure sensory' effects. Among these, there is a small but interesting literature on the critical flicker frequency (*CFF*).[6] Von Shiller, the first to describe the influence of heteromodal input upon the *CFF*, reported facilitation of fusion in a number of *S*s with the concomitant presentation of consonant tones, but enhanced flicker with dissonant combinations. These results he related to a phenomenal dimension of roughness, assumed to be common to both modalities. More recently, Ogilvie also found a higher *CFF* in the presence of in-phase flutter than in the presence of out-of-phase flutter. Continuous noise and stimulus-luminance, meanwhile, were found to have no effect on fusion. Kravkov, in contrast, had earlier reported the effect of a continuous tone on the *CFF*—raising it in the fovea and lowering it in the periphery—the maximal effect occuring at onset. He had further reported on increased *CFF* with a bright source and a lowered *CFF* with a weak source. Grignolo, Boles-Carenini, and Cerri have, however, claimed increases in *CFF* both for fovea and periphery with auditory stimulation from 2000 to 4000 ~ and 55 to 85 db.; Gorrell has reported lowering with a tone in the same frequency range, and McCroskey has obtained similar data with continuous noise. Finally, Kravkov has described different effects as a function of the color of the stimulus-light. Auditory stimulation was found to raise foveal *CFF* for a source of 630 mμ but to lower it for a source of 520 mμ. This difference he related generally to concomitant changes in excitability and excitation-level.

Heteromodal influences are conceptually obscure phenomena, but clarifying their significance for the present is less important than is explaining the seeming contradictions in the empirical data. Many of the reports are fragmentary and leave much to be desired in the matter of methodology. Accordingly, the present experiment was designed as a comprehensive evaluation of the effect of auditory input upon the *CFF* when several apparently important variables are treated parametrically and varied simultaneously.

Method. (1) Experimental design. The variables chosen for manipulation were frequency and intensity of auditory input and color of interrupted light. These were incorporated into a 3 × 3 × 3 design. The auditory frequencies were 290, 1050,

[6] P. von Shiller, Das optische Verschmelzen in seiner Abhängigkeit von heteromodaler Reizung, *Z. Psychol.*, 125, 1932, 249-264; Die Rauigkiet als intermodaler Erscheinung, *ibid.*, 127, 1932, 265-289; J. C. Ogilvie, Effect of auditory flutter on the visual critical flicker frequency, *Canad. J. Psychol.*, 10, 1956, 61-68; The interaction of auditory flutter and *CFF:* The effect of brightness, *ibid.*, 10, 1956, 207-210; Kravkov, Action des excitations auditives sur la frequence critique des papillotements lumineux, *Acta Ophthal.*, 13, 1935, 260-272; Critical frequency of flicker and indirect stimuli, *C. R. (Dak) Acad. Sci. URSS*, 22, 1939, 64-66; A. Grignolo, B. Boles-Carenini, and S. Cerri, Researches on the influence of acoustic stimulation upon the critical fusion frequency of light stimulation, *Rivista Oto-neuro-oftal.*, 29, 1954, 56-73; R. B. Gorrell, The effect of extraneous auditory stimulation on critical flicker frequency, Unpublished Doctoral dissertation, Clark University, 1953; R. L. McCroskey, Jr., A research note on the effect of noise upon flicker fusion frequency, USN School of Aviation Medicine, Joint Project Report, 70, 197; Kravkov, *Rabota organov chuvstv,* 1949.

and 3900 ~; the loudnesses were 0, 40, and 80 phons; and the dominant wave-lengths of the three color-sources were 490.5, 538.0, and 605.7 mμ. Three groups of 12 *S*s each were used, each group being exposed to all combinations of auditory stimulation while making judgments for one color of light. An over-all triple-classification analysis of variance was performed on the data of all 36 *S*s, then additional analyses were carried out separately on each of the three color-groups.

Since the *S*s were arbitrarily assigned to groups as they appeared for testing, a check on group-comparability was obtained by evaluating *CFF*s for white light obtained after the color-series A simple analysis revealed no significant differences among group means [$F = 3.10$, *df.* (2, 33), $P > 0.05$]. In addition, comparison of the three groups for the colored-light-no-sound condition failed to yield a significant *F*-ratio [$F = 0.12$, *df.* (2, 33), $P > 0.05$].[7]

(2) Subjects. The 36 *S*s were college men from 18–28 yr. of age, with no known visual or auditory defects. All were requested to refrain from excessive physical exercise or the use of any drugs for at least 6 hr. prior to testing.

(3) Apparatus. Intermittent stimulation was produced by interposing an episcotister between the light-source and *S*. Constant speeds of rotation were achieved with an Eaton Dynamatic motor operated through a Sola voltage-regulator and controlled by a 3600° Helipot Potentiometer. A conversion-scale of potentiometer readings to exposures-per-second was constructed with the aid of a GRC Type 631-BL Strobatic. The light-dark ratio of the disk was 1 : 1. A 20-min. warm-up was allowed the motor before each day's testing.

The light-source consisted of two 150-w. Sylvania C7-A filament-lamps housed in a diffusion-box and operated at 113 v. to provide a color temperature of 2854°K (Illuminant *A*). (Testing was completed well before any appreciable loss in the efficiency of the bulbs occurred.) A metal mask with a single circular aperture mounted in the forward end of the box provided a test-patch of 1°. Directly in front of the mask were mounted an Eastman Kodak two-log unit neutral-density wedge, balancing wedge, and the appropriate color-filter. The latter was one of three Eastman Kodak Wratten narrow-band filters: 72B, 74, or 75. All were high and approximately equal in excitation-purity. Illumination-intensity at the eye was adjusted to 0.25 ft.-candles for all colors. A Bausch and Lomb chin-rest insured the proper viewing position for all *S*s. A telegraph-key connected to a small lamp at *E*'s station allowed *S* to report thresholds.

Auditory input was provided by means of a Hewlett-Packard Model 200 AB Audio-Oscillator, controlled through a Hewlett-Packard Model 350A attenuation pad, and delievered to *S* through Permaflux PDRIO earphones. A Wichita Electronic Switch was used to control rise-time and eliminate clicks. The impedances of earphones and oscillator were matched with the aid of a Thordarson T-65594 transformer.

(4) Procedure. All *CFF*s reported in the present study are monocular *CFF*s for the fovea. Each *S* was seated in the test-room, told to cover his left eye with an eye-patch provided him, and then shown how to sit for viewing, and how to operate the signal-key. The transitions from flickering-to-steady light and vice versa were demonstrated for him and he was asked to signal whenever he became aware of such changes. He was next allowed three flickering-to-steady and three steady-

[7] The monochromatic test-sources were matched for brightness.

to-flickering practice-trials. The earphones were then fitted to his head, and he was told that on certain trials he would hear a tone but that he was to continue reporting visual changes regardless of its presence or absence. Opportunity for questions about procedure was allowed as a part of a final review of method, and then *S* was allowed 10 min. of dark-adaptation. Testing began with a signal of three taps on the key by *S*.

Each *S* made judgments on 45 trials. Each trial consisted of one flickering-to-steady and one steady-to-flickering determination. The 45 trials represent the total of 5 trials each of 9 combinations of loudness and frequency randomized in a different order of presentation for each of the 12 members of each color group. On the trials when tones were presented, auditory-input was continuous for the interval

TABLE I

CFF FOR THE SEVERAL COMBINATIONS OF AUDITORY FREQUENCY, LOUDNESS-LEVEL, AND DOMINANT WAVE-LENGTH

Cell-entries are means based on groups of 12 *S*s serving under one of the three conditions of visual stimulation

Frequency	Wave-length	Loudness 0	40	80	Mean
290	605.7	20.01	19.62	19.45	
	538.0	20.89	20.82	20.82	20.52
	490.5	20.65	20.89	21.52	
1050	605.7	19.91	19.39	19.41	
	538.0	21.14	20.90	20.60	20.49
	490.5	20.72	21.07	21.23	
3900	605.7	19.79	19.64	19.28	
	538.0	20.95	20.82	20.84	20.50
	490.5	20.59	21.04	21.55	
Mean		20.52	20.47	20.52	

between presentation of light and the signaling of change. After the color-trials had been run, 5 additional trials were conducted with white light and no sound.

Results: (1) Data. Table I presents averages for the several conditions of judgment, Table II summarizes the over-all analysis of variance, and Fig. 1 shows the interaction between loudness-level and the color of the test-source. An examination of these results yields several impressions: (a) that auditory input does exert an influence on *CFF;* (b) that this effect is rather small in magnitude and rather complex, and (c) that, in specific details, the present findings do not appear to be consistent with others previously reported.

Although the literature on intersensory effects yields the impression of dramatic changes, close examination of the papers on *CFF* indicates the maximal shift reported to be about 10% (McCroskey).[8] Meanwhile, the 2–

[8] McCroskey, *op. cit.*, 197.

TABLE II

SUMMARY OF ANALYSIS OF VARIANCE PERFORMED ON *CFF*-DATA FOR ALL CONDITIONS AND ALL *S*s

Source*	*SS*	*df.*	*MS*	*F*	*P*
Total	9,644.15	323			
Between *S*s	9,352.84	35	267.22	264.57	.01
Between wave-lengths (blocks)	129.66	2	64.83	0.23	NS
Error (between)	9,223.18	33	279.49		
Within *S*s	291.31	288	1.01		
Frequency (columns)	0.06	2	0.03	0.16	NS
Loudness (rows)	0.21	2	0.10	0.67	NS
Frequency × loudness	1.06	4	0.26	0.35	NS
Frequency × wave-length	0.38	4	0.10	0.67	NS
Loudness × wave-length × color	16.88	4	4.22	28.13	.01
Frequency × loudness × wave-length	143.06	8	17.88	24.16	.01
Error (within)	129.66	264	0.49		
Error_1 (within)	12.59	66	0.19		
Error_2 (within)	19.30	66	0.15		
Error_3 (within)	97.77	132	0.74		

* The various tests of significance (E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 296) were made as follows: frequency and frequency × wave-length against error_1 (within), loudness and loudness × wave-length against error_2 (within), loudness × frequency and frequency × loudness × wave-length against error_3 (within), and wave-length against error (between). The error (within) term was used in obtaining error_3 (within).

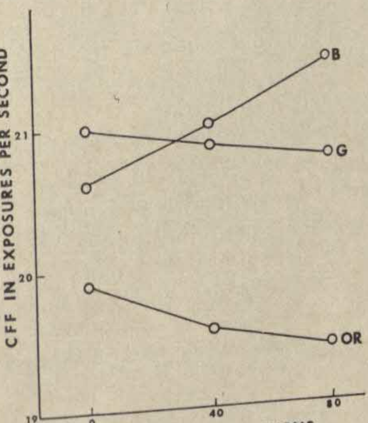


FIG. 1. MONOCULAR *CFF*, AS A FUNCTION OF COLOR OF TEST-SOURCE AND LOUDNESS OF AUXILIARY AUDITORY INPUT (*OR* = orange-red, *G* = green, *B* = blue.)

4% changes of the present study, while comparable to those found by Ogilvie, are clearly less than those reported by Grignolo, *et al.* and by Kravkov.[9] The complexity of the effect is indicated by the fact that while none of the main sources of variance, except individuals, proved significant, one of the three first-order interactions and the second-order interaction were clearly reliable. The inconsistencies elaborated below are difficult to explain. The several previous experiments differ widely in methodological particulars, and none represents a comprehensive evaluation of the role either of the inter- or the intra-modal variables. They vary, moreover, so greatly in completeness of report, that comparisons only of the most general sort can be made.

(2) The role of color. As was indicated above, color alone, when equated for intensity at the eye, had no effect on the *CFF*.[10] When combined with auditory input, it became a significant variable. Fig. 1 indicates that the addition of sound resulted in an increase in *CFF* for the blue source, no change for the green source, and a decrease for the orange-red source. Allen and Schwartz, meanwhile, have reported increases both with short and with long wave-lengths, and Kravkov has claimed a decrease for green and an increase with orange-red.[11] These differences are, for the moment, unreconcilable. They cannot be attributed either to the frequency or to the loudness-level of the sound, for Kravkov worked within the range employed in the present experiment on both dimensions. One possible source of difference, although Kravkov does not provide the necessary information on this point, is in the intensity of the test-light. In the same paper in which he describes the color-differences, he reports that sound added to a dim green test-source resulted in a decrease in the *CFF*, while sound added to a bright source resulted in an increase.

(3) The effect of loudness-level on the CFF. Fig. 1 indicates that the *CFF*-by-loudness functions for orange-red and blue are monotonic. This impression is supported by appropriately significant between-levels $F =$ ratios in the variance-analyses performed on the data for each color group (orange-red, $F = 21.92$, *df.* (2, 22), $P < 0.01$; green, $F = 2.84$, *df.* (2, 22), $P > 0.05$; blue, $F = 5.75$, *df.* (2, 22), $P < 0.01$). This finding

---

[9] Ogilvie, *op. cit., J. Psychol.,* 10, 1956, 207-210; Grignolo, Boles-Carenini, and Cerri, *op. cit.,* 56-73; Kravkov, *op. cit., C.R. (Dak) Acad. Sci. URSS,* 22, 1939, 64-66.

[10] Selig Hecht and S. Shlaer (Intermittent stimulation by light. V. The relation between intensity and critical frequency for different parts of the spectrum, *J. gen. Physiol.,* 19, 1936, 965-977) report differences in *CFF* as a function of wave-length, but they used a 19° patch, and there is no indication that they controlled intensity at the eye.

[11] F. Allen and M. Schwartz, The effect of stimulation of the senses of vision, hearing, taste, and smell upon the sensitivity of the organs of vision, *J. gen. Physiol.,* 24, 1940, 105-121; Kravkov, *ibid.,* 64-66.

agrees with Kravkov, who found a change in the *CFF* as auditory intensity was increased from 25 to 95 db., with the important exception that the direction of the effects is reversed.

(4) The effect of auditory frequency on the CFF. Pitch as a main variable failed to exert any influence upon the *CFF*. Nor did it have an effect when paired either with color or with loudness-level. When co-varied with both, however, it participated in a significant triple interaction (Table II). This outcome would appear to be somewhat at variance with Gorrell who reported that a high-frequency tone lowered the achromatic *CFF*, while a low tone had no effect. At variance with both, is Grignolo's claim that tones from 2000 to 4000 ~ raised achromatic *CFF*.

It is presently extremely difficult to comment on the current results in physiological terms. While certain findings—such as differences related to strong and weak magnitudes either of primary or of auxiliary stimuli and differences associated with foveal and peripheral *CFF*, may be discussed, at least generally, in terms of the mechanisms of neuronanatomical convergence,[12] this approach does not readily accommodate the data on the role of color or allow one to anticipate the nature of the color-by-loudness interaction. Meanwhile, the significant color-by-loudness and color-by-loudness-by-pitch interactions may have interesting implications for Steven's dichotomy of prothetic and metathetic sensory dimensions and his assumption of fundamentally different underlying mechanisms.[13]

Summary. Thirty-six Ss were employed in a 3 × 3 × 3 design to study the influence of auditory input upon monocular, foveal *CFF* for lights of different color. Three groups of 12 *Ss* each were tested with one of three dominant wave-lengths. 490.5, 538.0, and 605.7 mμ. Each group experienced all combinations of three loudness-levels (0, 40, and 80 phons) and three frequencies (290, 1050, and 3900 ~). The method of limits was used, with threshold approached from both directions. An analysis of variance revealed complex intersenory relationships. When the test-source was orange-red, the *CFF* decreased with an increase in the loudness; when it was green, no change was effected; and when it was blue, the *CFF* increased. Frequency influenced the *CFF* only when acting in combination with test-source color and loudness-level. Comparisons were made between these results and earlier findings.

[12] See, for example, J. D. Harris, Some relations between vision and audition, 1950, 41-46. The matter of the locus of interaction poses many problems. Meanwhile Grignolo, *et al.* (*op. cit.*, 56-73) suggest centrifugal excitation from the mid brain to the retina as a plausible explanation. R. Granit (*Receptors and Sensory Perception*, 1955, 104-110) presents evidence for such excitation.

[13] S. S. Stevens, On the psychophysical law, *Psychol. Rev.*, 64, 1957, 153-181.

THE UNIQUENESS OF PATTERNS OF SKIN-CONDUCTANCE

By J. W. KLING and HAROLD SCHLOSBERG, Brown University

Skin-conductance has long been recognized as a useful index of the level of general tension or activation. Conductance-levels are high when *S* is alert, interested, or anxious; and they fall as *S* becomes drowsy, or adapts to the situation, or become bored. Superimposed on the level of conductance are the relatively transient changes (*GSR*). It is important to distinguish between these two indices: the *GSR* is part of S's response to stimuli of the moment, while conductance-level seems to reflect *S*'s position on a continuum from low to high activation.[1] Indeed, it has recently been suggested that conductance-level may prove useful as an objective indicator of motivation.[2]

In a typical experiment, a stimulus-object (word, shock, or picture) is presented, and the change in skin-resistance or conductance over the next few seconds is observed. Then, after a suitable pause to allow the basal level of conductance to be approached, a new stimulus is presented. The *magnitude* of the change in conductance is utilized as the datum in most studies, but it is possible that the *pattern* of change would reveal valuable additional information. If conductance-readings are taken more or less continuously during a session, not only the magnitude of the transient response may be obtained, but the time-course of the approach to the resting level of conductance may also be noted. It seems likely that such patterning of conductance-readings is related to certain characteristic response-tendencies of *S*, and to the extent that this is true one should expect successive runs through the same procedure to produce records from a given individual which have a greater-than-chance degree of similarity.

To ascertain the extent to which a given *S* showed consistency in reacting to a laboratory-situation, a comparison was made of the patterns of skin-conductance obtained on two separate occasions. The conductance-values were recorded during two 5-min. rest-pauses which were interpolated between successive blocks of practice in rotary-pursuit. All *S*s were receiving

* Received for publication October 29, 1959. The pursuit-scores and group conductance-values were reported in a previous paper (J. W. Kling, J. P. Williams, and Harold Schlosberg. Patterns of skin-conductance during rotary-pursuit, *Percept. Mot. Skills*, 9, 1959, 303-312). This was one of a series of studies supported in part by Thomas Lipton, Inc. We wish to thank Joanna P. Williams who gathered the first conductance-scores.

[1] Harold Schlosberg, Three dimensions of emotion, *Psychol. Rev.*, 61, 1954, 81-88.

[2] R. B. Malmo, Anxiety and behavioral arousal, *Psychol. Rev.*, 64, 1957, 276-287.

the same treatment during these pauses, and the manner in which they responded to the procedure was examined. If the changes in conductance were merely the reflections of an experimental procedure, all *Ss* should show approximately the same pattern. If, at the other extreme, the laboratory procedure was of little or no importance in determining the pattern of conductance-changes, there should be no noticeable similarity among the *Ss* receiving the same treatment. If, however, the pattern of change is primarily a function of how the procedure influences each *S*, there should be some general patterning, but the within-*S* similarity should be greater than the similarity between *Ss*. In other words, the pattern obtained from a given *S* on Rest 1 should be recognizably similar to that obtained on Rest 2; and these pairs of patterns should be distinguishable from the patterns obtained from other *Ss*.

Methods and procedure: (1) Apparatus. Readings of skin-conductance were obtained from silver-disk electrodes attached to the plantar surface of each foot. The disks were 1⅛ in. in diameter, of pure silver, and unchlorided. A non-irritating paste of kaolin, glycerin, and Ringer's solution insured adequate contact between the skin and the electrode. A sponge cushion backed the electrode, and Scotch tape was used to strap the unit to the foot. *S* was in series with a 1-v. source, a potential divider, and a microammeter. Readings thus were obtained directly in μΩ, the distribution of which appears to be essentially normal.[3] To minimize electrode-polarization, the circuit was turned off except during readings. Any remaining polarization was balanced out by quickly reversing the direction of current-flow during the reading, and recording the average of the two positions of the meter-pointer. Readings took approximately 2 sec. and were made every 15 sec.

(2) Subjects. Forty-eight women, undergraduate volunteers, were recruited from introductory sections in psychology for an experiment in "how we acquire skills." None had had previous experience in rotary-pursuit which she was given during blocks of work. All were reassured concerning the attachment of wires to their feet by promising that no shock would be used, but they were not told the purpose of the electrodes nor the exact nature of the experiment.

(3) Procedure. The *Ss* were assigned at random to one of three groups for practice. After *S* was introduced to the situation, the soles of her feet were cleansed and dried, and the electrodes attached. She was then seated upon a stool, which was adjusted for proper performance in rotary-pursuit. Her feet were placed upon a footrest which did not touch the electrodes or the nearby skin. She was told to keep her feet in that position at all times until the conclusion of the experiment. An easy-chair was placed next to the stool, and she was told that later she would be asked to slide from the stool into the chair for a rest-period, but that her feet were to be kept on the footrest. The easy-chair was utilized in an attempt to facilitate relaxation during the rests.

After instructions in how to perform in rotary-pursuit, *S* received a 5-min. practice-period. The rotor was then stopped, and *S* was instructed to slide into the easy-

[3] Harold Schlosberg and W. C. Stanley, A simple test of the normality of twenty-four distributions of electrical skin conductance, *Science*, 117, 1953, 35-37.

chair. After 4½ min. of rest, *E* announced that it would soon be time to resume work, "so get ready to move when I tell you." Then the order to move to the stool was given, and, after 5 min., the rotor was started and *S* was told to resume pursuit.

Following the next block (either 10 or 15 min.) of pursuit-practice, the rotor was again turned off and *S* instructed to return to the easy-chair. The procedure of Rest 1 was duplicated as closely as possible. Throughout both periods, conductance-readings were obtained at 15-sec. intervals. There were thus 20 readings for each rest, the last two in each case coming after a warning, and then a request to move back to the stool.

Results: (1) *Pairings by judges.* The conductance-scores for the last minute of practice, the 5-min. rest, and the first minute of post-rest practice were plotted for each *S* for each rest. Since the three groups differed in the conditions of work (*i.e.* degree of massing of practice and length of work-blocks), each group of 16 *S*s was treated separately. The 16 curves depicting the patterns obtained during Rest 1 for a given group were mounted on a large sheet. The 16 patterns obtained during Rest 2 were plotted on individual sheets. The task of the judges was to place one of the curves of Rest 2 with the curve of Rest 1 which seemed to be most similar. To forestall preconceived notions of how the curves should be paired, a judge was told merely that each of the curves on the large sheet had a related curve in the pack he held, and the matching was to be done on the basis of any and all cues he could employ. Until they had completed the task, judges were not told that these were conductance-curves.

Fifteen staff members and graduate students served as judges. They worked as individuals, and were free to change the pairings until they were satisfied with their matching. The materials were then removed, and a new set of 16 pairs of curves was presented. The order of presenting the groups of curves to the judges was so balanced that any practice-effects were distributed equally among the three groups.

The task of the judge may be appreciated by covering the axis-labels of Fig. 1 and noting the degree of similarity of the Rest 2 curves with those of Rest 1. The *S*s represented in Fig. 1 were selected at random, and are not extreme in the degree to which they differ. In general, most judges found a few curves in each group which were very difficult to match with the sample (Rest 1) curves, but the majority of matches seemed fairly easy to make.

The results of the judgments support the hypothesis that there are observable similarities within pairs of curves, but do not indicate what cues the judges were using. Judges reported that they based their judgments on both the form and the level of the curves. The success of the judges in carrying out the matching is indicated in Table I, which summarizes the results of this procedure.

(2) Objective comparisons. The use of blind matching by judges has one advantage over objective comparison: The judges may pick up cues that would be missed in direct measurements of such complex curves. The

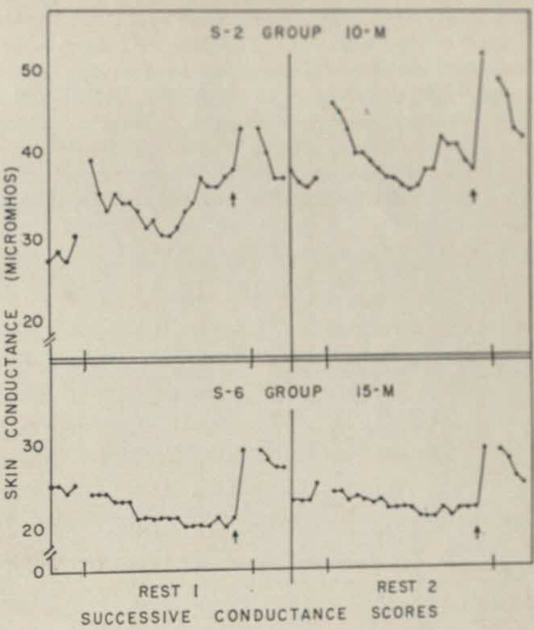


FIG. 1. TWO EXAMPLES OF CONDUCTANCE-PATTERNING DURING REST-PAUSES
The first four points were obtained during the last minute of work. The curves are broken to indicate the start and the end of the rest-periods. Arrows indicate the warning that the end of the rest-period is approaching. The last four points show conductance during the first minute of post-rest work. Throughout, conductance-readings were taken at 15-sec. intervals.

judges, however, clearly varied in skill, at least partly because of obvious differences in interest in the task, and it seemed advisable, therefore, to attempt some objective measure of the similarity.

To eliminate absolute level of conductance, which tended to increase between the first and second rest-periods, each curve was adjusted to have a mean conductance of 50 μΩ. This was done by summing the 20 readings that constituted one record, multiplying the reciprocal of the sum by 1,000, and then using the resulting conversion-factor to replot the curve. From these adjusted data, a discrepancy-score was

TABLE I

MATCHES OF CONDUCTANCE-CURVES OF REST 1 AND REST 2

Group	Correct matches			
	N	Mean	Range	p of mean*
10-S	16	7.3	3–12	0.00008
10-M	16	5.5	1– 9	0.004
15-M	16	6.0	1– 9	0.0006

* Estimates of probability of obtaining by chance the mean (or greater) number of correct matches, based on Poisson limits.

computed for each pair of curves—the sum of the absolute differences between paired points. The resulting distribution of 48 discrepancy-scores was then compared with a similar distribution of scores for random pairs of records; *i.e.* each curve for Rest 1 was compared with that of a randomly selected *S* for Rest 2, and chance-difference scores were similarly computed.

A simple estimate of the reliability of the differences between the two sets of scores was obtained with Wilcoxon's test for paired replicates. For Group 10-*S*, the difference was significant at the 2% level, and, for the other two groups, at the 1% level. There is a significant tendency for the individual to be less variable from rest to rest than might be expected on the basis of random variability in patterns of conductance. Not only can judges recognize pairs of patterns which belong together because they came from the same *S*; but when the factor of amplitude is removed and mere departure from curve form is utilized for objective comparisons, a greater-than-chance consistency in patterning remains.

Discussion. Measures of skin-conductance at 15-sec. intervals during rest-pauses reveal patterns of change which are fairly similar within individuals. The degree of correspondence of the two patterns for each person is not perfect, and there are a few individuals who produced apparently different patterns on the two occasions, yet the similarity within individuals is striking when the conditions of the experiment are considered. Although the salient features of the rests (start, warning of end, and command to return to work) were determined by a fixed time-schedule, the content of the intervals was casual conversation. Thus, one might anticipate that under conditions of exact duplication of procedure, even greater correspondence among the pattern of conductance might be observed.

Basically, the patterns seem to be generated by *S*'s responses to the change in procedure (start of rest, end of rest). Although the manner in which an *S* responded was, in most cases, highly individualistic, certain typical patterns do emerge from an over-all consideration of the curves. The most frequently seen form (in 13 of 48 cases) is one which falls rapidly from the peak at the start of rest, reaches a low point, and starts to rise prior to the occurrence of the warning to return to work, *e.g.* *S*-2, Fig. 1. Although differing in the amount of change and the point at which the low is reached, the *S*s used in this study show a pattern that suggests a compensatory rise in conductance after the initial period of 're-laxation.' It should be stressed that no special signals were provided by the experimental procedure at the time this reversal of trend occurred, and, to the extent that conductance is a valid index of general tension, the scores suggest that these *S*s may be responding with increased tension to cues produced by their own relaxation.

Other patterns easily recognized include the rapid fall, at the start of rest, to a low which is maintained until the warning that work is to resume (eight cases); and the continual drop throughout the rest until the warning (six cases). Finally, there are 10 cases in which the pattern during the rest-period is essentially flat until the warning, at which time conductance rises. Eleven *Ss* produced patterns not easily categorized or described, yet in most cases the correspondence for each *S* from Rest 1 to Rest 2 is recognizable.

If all *Ss* showing a certain pattern are brought together, there are still differences in the range of conductance-values, the point in time at which a low value is reached, and the like; but, because of the similarities, the matching scores of the judges and the measures of consistency given by the discrepancy-scores are actually conservative estimates of the reliability of response within *Ss*. Where judges erred in matching a Rest 2 pattern with a Rest 1 pattern, the errors frequently were 'good' ones. Some patterns can be well matched by a curve from another *S* because both show similar changes, yet the consistency of both persons involved in the 'mismatch' may be high. Similarly, the random matching procedure employed in lieu of a known distribution of discrepancy-scores sometimes matched two persons who had produced similar conductance-patterns. Thus, neither of these techniques is completely appropriate to the demonstration of the true degree of within-*Ss* consistency, but they were employed since the present data were not amenable to the usual correlational techniques.

The consistency of an individual's level and pattern of skin-conductance suggests several lines of investigation. In the first place, it would be important to determine how stable this pattern is under varying experimental conditions. Furthermore, it would seem profitable to compare the level and form of these curves with other response-characteristics of the *Ss*. In general, the present results suggest that the amplitude and latency of the sharp initial rise in conductance (GSR) is no more important than the subsequent recovery-period in estimating the effect of (and adaptation to) a variety of experimental situations.

Summary. The skin-conductance of 48 women was followed through two rest-periods interpolated between sessions of practice on a pursuit-rotor. The level and pattern of the individual records showed considerable stability for each *S*; not only did judges have considerable success in pairing the corresponding records for each rest-period, but quantitative analysis of the patterns also showed better-than-chance similarity within *Ss*. The results suggest that conductance-patterns might prove useful in the study of individual differences in responsiveness to various situations.

THE INFLUENCE OF SYNTACTICAL STRUCTURE ON LEARNING

By WILLIAM EPSTEIN, University of Kansas

This experiment was designed to study the role of syntactical structure in verbal learning. Syntax can be defined as the generalized pattern or schema which is imposed upon the reservoir of available words and determines the sequences of these words. The fact that verbal messages in ordinary usage are encoded according to a set of grammatical rules may make the learning of natural linguistic units very different from the learning of a series composed of independent items, *e.g.* a list of nonsense-syllables.

Sequential or transitional-probability orderings and syntactical arrangements must be distinguished at the outset. A sentence usually entails a high degree of transitional probability among its components. Nonetheless, a sentence cannot be defined simply as a highly probable sequence of words. The distinction is clearly expressed by Deese,[1] and also by Chomsky. Concerning this point, Chomsky writes: "If we rank the sequences of a given length in order of statistical approximation to English, we will find both grammatical and ungrammatical sequences scattered throughout the list; there appears to be no particular relation between order of approximation and grammaticalness."[2]

The ability to produce and to recognize grammatical utterances, therefor, is not based on any notions of transitional probability. Nor does the experience of syntactical structure depend on semantic meaningfulness or familiarity. Grammatical structure can be recognized easily even in an arrangement of nonsense-syllables.[3] It should be possible, therefore, to study the influence of syntactical structure independently of meaningfulness, familiarity, or sequential probability.

Some evidence bearing on this factor is cited by Osgood, one of whose students "compared the ease of learning nonsense sequences that retained the structure of the English sentences from which they were derived, for

* Received for publication December 28, 1959. This study was supported by Grant M-3600 from the United States Public Health Service.

[1] James Deese, *The Psychology of Learning,* 2nd ed., 1958, 331.

[2] Noam Chomsky, *Syntactic Structures,* 1957, 17.

[3] I do not mean to imply that the listener or reader must be aware of the grammatical aspects of the message to respond to the information provided by syntax.

example, *The maff vlems oothly um the glox nerfs,* with matched materials in which the grammatical cues had been eliminated, for example, *maff vlem ooth um glox nerf*"[4] The structured series were learned more easily than random lists. Further evidence is provided by the present investigation in which the learning of structured material and matched, unstructured material was compared.

Materials. Six categories of material were composed, each consisting of two 'sentences.' Category I contains two sentences composed of nonsense-syllables in combination with two functional words which have no referential meaning, *e.g.* articles, conjunctions, and prepositions. Grammatical tags, such as *ed* on past tense verbs and *s* on plural nouns, were appended to the syllable-stems to simulate the requirements of English syntax. The result is a series of nonsense-syllables which is readily perceived to be grammatically structured. In Category II, these syllable-stems are presented again in the same order but without the appended grammatical tags. Category III repeats the material contained in Category I, but the items are arranged in random order. As a consequence of the omission of the tags in Category II and the randomization of items in Category III, the four sentences comprising these two categories are devoid of any discernible syntactical structuring.

In Category IV, the identity of the original syllable-stems and the order with which they are presented in Category I are retained. The only difference between the two categories is in the position of the appended grammatical endings. In the fourth category, the positions of the tags have been shifted in an effort to induce the formation of a competing syntactical pattern, that is, a structure which is not congruent with customary English usage.

For purposes of comparison, the remaining two categories consist of meaningful words. These words are arranged in an order which is *sententially* meaningless and within which there exists a very low level of transitional probability between neighboring words. Category V contains two series of words so ordered as to meet the demands of syntactical structure. Category VI contains the same words in random, unstructured order.

The sentences were typed separately on 5 × 8-in. index cards in a horizontal line, as is customary for written material in English. To enhance the grammatical character of the material in Categories I, IV, and V, the first word in each of these sentences was capitalized and each sentence was closed with a period. In the remaining three categories both the capital letter and the period were omitted. The material is reproduced in Table I.

Subjects. *Ss* were 192 students in an introductory course in psychology. None had prior experience with verbal, learning tasks in an experimental situation. The *Ss* were randomly assigned to six equal groups, each group learning the material in one of the six categories.

---

[4] C. E. Osgood, A behavioristic analysis of perception and language as cognitive phenomena, in J. S. Bruner, *et al.*, *Contemporary Approaches to Cognition,* 1957, 88. I am indebted to Dr. Osgood for making available an unpublished report of this work, by Swanson, which together with his own comments, facilitated the present investigation.

Procedure. Ss performed the required tasks individually. The following instructions were read to each:

This is an experiment in verbal learning. I am going to show you a series of nonsense-syllables (words) for 7 sec. Try to learn all the syllables (words) in the series in the order in which they are arranged. Distribute your attention evenly among the items so that you can learn all of them. When the 7 sec. are over you will be given 30 sec. in which to write down what you remember. If your response is incorrect in any way, I will show you the series again. We will repeat this process until you reproduce the series perfectly. It will be very helpful to me if you guess or fill in dashes where you are not certain of the item.

After the instructions were concluded, one of the sentences in the assigned category was presented to *S* for 7 sec. Half the *S*s in each group learned Sentence 1

TABLE I

CONTENT OF THE SIX CATEGORIES

Category	Sentence
I	(1) A vapy koobs desaked the citar molently um glox nerfs.
	(2) The yigs wur vumly rixing hum in jegest miv.
II	(1) a vap koob desak the citar molent um glox nerf
	(2) the yig wur vum rix hum in jeg miv
III	(1) koobs vapy the um glox citar nerfs a molently
	(2) yigs rixing wur miv hum vumly the in jegest
IV	(1) A vapy koobed desaks the citar molents um glox nerfly.
	(2) The yigly wur vums rixest hum in jeging miv.
V	(1) Cruel tables sang falling circles to empty bitter pencils.
	(2) Lazy paper stumbled to shallow trees loudly from days.
VI	(1) sang tables bitter empty cruel to circles pencils falling
	(2) loudly trees paper from days lazy shallow to stumbled

first, and the other half learned Sentence 2 first (see Table I). Immediately after the end of this learning period, *S* was given 30 sec. in which to reproduce in writing all that he could remember of the sentence. The record then was checked, and, if it was incorrect in any way, the sentence was presented again for a second 7-sec. trial, followed by another test. This procedure was repeated until a perfect reproduction was obtained. To be scored as perfect, a reproduction had to include all of the items in the correct order. Errors in spelling were overlooked if they did not affect the pronunciation of the item. When the criterion was reached with the first sentence, the second sentence was presented and learned under the same conditions.

Results. The main results for the six categories, in terms of mean trials to criterion for the pair of sentences, are as follows: I, 5.77 (SD = 2.39); II, 7.56 (SD = 3.42); III, 8.15 (SD = 3.16); IV, 6.90 (SD = 2.29); V, 3.50 (SD = 1.58); VI, 5.94 (SD = 2.25). An analysis of variance of the results yields an F-ratio of 13.35 with 5 and 186 *d.f.*, which is highly significant ($p < 0.001$).

Various comparisons between the categories are possible. Some of the more interesting may be noted. If syntactical structure facilitates learning, then Category I should require fewer trials than Categories II, III, or IV, and Category V should require fewer trials than Category VI. In a more conjectural vein, the interference produced by dissonant structure might be expected to make Category IV the most difficult. A comparison of Categories I and VI gives an opportunity to assess the strength of the syntactical factor relative to meaning and familiarity. I had no specific expectation concerning this comparison.

The *q*-statistic for comparing individual means in the one-variable case was used to determine the significance of these differences.[5] Of the eight differences tested, the following four were significant at the 5% level: I *vs.* II, I *vs.* III, I *vs.* V, and V *vs.* VI.

Discussion. The results clearly indicate that syntactical structure facilitates learning. The only comparison which did not show the effect was I *vs.* IV, although the obtained difference of 1.13 trials was in the expected direction. Two factors may account for the absence of a significant difference here: (1) In both categories, Sentence 1 begins with the same two items. (2) The presence of tags in Category IV may have encouraged *S* to restructure the material into a syntactically congruent unit, thereby improving performance. The absence of a significant difference between the first and sixth categories also is of considerable interest. It suggests that the facilitating effects of syntax in Category I compensated for the advantages provided by meaning and familiarity in Category VI.

Several different explanations of the present results are possible. I shall simply present them here, with no attempt to evaluate their relative merits.

(1) On the basis of Osgood's analysis of language, the facilitation of learning in our experiment might be understood as another demonstration of the operation of "predictive integrations in the grammatical mechanisms that interrelate larger message events."[6] The frequency with which grammatical redundancies occur in ordinary language is believed to result in strong predictive integrations in the nervous system that match the structure of the language. These integrative systems make encoding and decoding of congruent messages easier by restricting the number of alternative responses and by ordering the probabilities attached to the various alternatives. The role of grammatical redundancies in the determination of response-selection and learning, according to Osgood, is demonstrated in

[5] W. J. Dixon and F. J. Massey, *Introduction to Statistical Analysis*, 2nd ed., 1957, 156-159.

[6] Osgood, *op. cit.*, 85.

Taylor's work using the 'cloze' procedure and also in Miller and Selfridge's work on contextual constraint.[7]

(2) An interesting alternative can be derived from Miller's analysis of the immediate-memory span.[8] Miller suggests that the size of the immediate-memory span is not determined by the amount of information per item in the testing material, *e.g.* words or numerals, but that people remember a constant number of 'chunks' of information irrespective of the amount of information in each. The learning of new material proceeds on the basis of the formation of chunks or the reorganization of the material into a small enough number of chunks, a process called *recoding.* When we can recode a sentence into that number of chunks which corresponds with the size of the immediate-memory span, we can recall the entire sentence. Material which is not syntactically structured may be harder to learn than structured material because the latter is already organized whereas the former can be organized into more efficient chunks only through the intentional efforts of the learner. There are surprising similarities between this explanation and one which could be derived from Köhler's view of the role of organization in learning.[9]

(3) Another possibility, not entirely unrelated to the preceding viewpoint, is that messages cast in syntactical form are learned more easily because different strategies of learning are employed for organized and unorganized material. When *S* is confronted with a random sequence of items, he may be led to distribute his attention unevenly and unsystematically among the items. Conversely, the presence of patterning may encourage a more systematic approach leading to more rapid learning. Some indirect evidence for the existence of different strategies can be derived from a recent investigation by Deese and Kaufman of serial effects in the immediate free recall of unorganized and sequentially organized verbal material.[10] The authors report that the order of recall for randomly arranged words correlated with the frequency with which the individual words were recalled. For organized textual material, however, the order of recall was correlated with the order of presentation of items in the list. In other words, textual passages were recalled roughly in their order of presenta-

---

[7] William Taylor, Cloze procedure: A new tool for measuring readability, *Journalism Quart.,* 30, 1953, 415-433; G. A. Miller and J. A. Selfridge, Verbal context and the recall of meaningful material, this JOURNAL, 63, 1950, 176-186.

[8] Miller, Information theory and memory, *Sci. Amer.,* 195, 1956, 42-46; Miller, The magical number seven; Plus or minus two: Some limits on our capacity for processing information, *Psychol. Rev.,* 63, 1956, 81-97.

[9] Wolfgang Köhler, *Gestalt Psychology,* rev. ed., 1947, 248-278.

[10] James Deese and R. A. Kaufman, Serial effects in recall of unorganized and sequentially organized verbal material, *J. exp. Psychol.,* 54, 1957, 180-187.

tion while for unorganized material "the items were emitted in a kind of primitive order of strength during test of recall."[11] It seems not unlikely that this reflects a difference in the way *S* attempts to learn and recall structured and unstructured material. Similar differences in approach may account for the present findings.[12]

Summary. An experiment was performed which showed that syntactical structure facilitates verbal learning apart from the contributions of meaningfulness, familiarity, and sequential probability. Possible explanations of this effect were discussed.

---

[11] *Ibid.*, 185.

[12] Serial-position curves prepared from our data did not differ in the manner to be expected from this viewpoint. The items in all the categories were learned in their order of presentation.

DISCRIMINATIVE PERFORMANCE OF MONKEYS IRRADIATED WITH X-RAYS

By ROGER T. DAVIS, University of South Dakota

Monkeys given 'whole body radiations' (*WBR*) have, according to recent reports, a narrower scope of attention than those not irradiated.[1] The generality of this finding is tested in the present experiment by a procedure that varies the number of objects presented for discrimination. Every monkey was given 140 problems, each requiring the discrimination of two objects: one rewarded and the other not rewarded with food. After 10 trials with one pair of stimulus-objects, a new pair was presented for the next problem. Three objects, superfluous to the solution of the problem, were introduced in Trial 2 and on the following even-numbered trials on each problem. If irradiated *S*s are less able to deal with extraneous cues, they should err relatively more often on even- than on odd-numbered trials.

Method. A pair of stimulus-objects which differed in many dimensions from each other, was presented 10 times to *S*. One member of the pair, designated 'A,' is accompanied by a reward (a raisin) when chosen by *S*, the other, designated by 'B,' is not rewarded. The positions (right and left) of A and B were varied according to a predetermined sequence. Every *S* was given 10 trials on each of 5 problems a day for 28 days. Seven hundred stimulus-objects (140 pairs of A- and B- objects and 420 neutral (*N*) objects) were required for this study, as an A- and B- object and three neutral (*N*) objects, five in all, were required for a single problem. On the odd-numbered trials, the A- and B- objects alone were presented; on the even-numbered trials, the three *N*- objects were added.

Thirteen rhesus monkeys were used in this investigation. Seven had previously received 1100 *r WBR* with X-rays in three doses, spaced a year apart. The remaining six *S*s had not been irradiated and served as a control group. All the *S*s had had extensive experience with problems requiring the discrimination of objects.[2] More

* Received for publication November 7, 1958. This investigation was supported in part by Grant M-530 (C-5) from the National Institutes of Health.

[1] A. A. McDowell, Comparisons of distractibility in irradiated and nonirradiated monkeys, *J. genet. Psychol.,* 93, 1958, 63-72; J. E. Overall and W. L. Brown, Narrowing of attention as a chronic effect of sublethal radiation. *USAF School of Aviation Medicine Report* No. 27, 1958.

[2] R. T. Davis, A. A. McDowell, C. W. Deter, and J. P. Steele, Performance of rhesus monkeys on selected laboratory tasks presented before and after a large single dose of whole-body X-radiation, *J. comp. physiol. Psychol.,* 49, 1956, 20-26; N. W. Heimstra, R. T. Davis, and J. P. Steele, Effects of various food-deprivation schedules on the discrimination learning performance of monkeys irradiated with X-ray irradiation, *J. Psychol.* 44, 1957, 271-281.

recently, however, they had been given a serial-learning problem which may have interfered to some extent with their ability to discriminate.[3]

Results. The effect of introducing additional unrewarded objects is presented graphically in Fig. 1. The left-hand side of the figure shows the mean performance within a problem averaged over the first 70 problems, that is, for the first 14 days of the experiment. The solid lines represent the *Ss'* performance on the five trials in each problem on which five objects

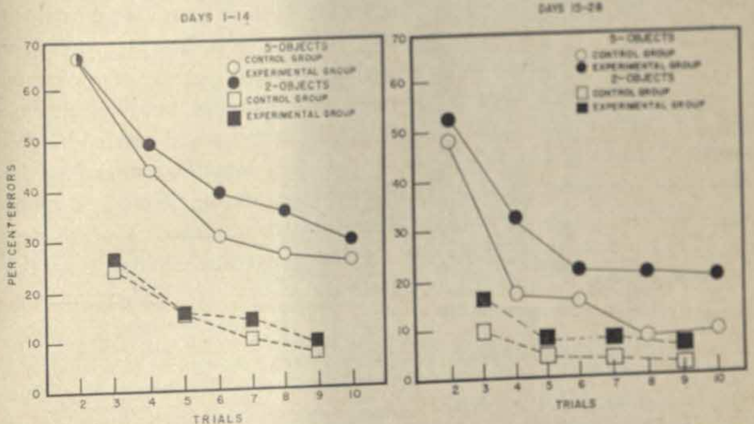


FIG. 1. INTRAPROBLEM PERFORMANCE OF MONKEYS IN THE DISCRIMINATION OF OBJECTS

The results on Days 1 to 14 and 15 to 28 are averaged trial-by-trial, and are shown as a function of prior irradiation and the number of stimulus-objects present on a given trial.

were employed. The dotted line shows the performance on trials with two objects. The filled circles and squares represent performance of the *Ss* in the experimental irradiated group and the open figures the performance of the control group.

An *S*'s performance on Trial 1 is determined by its chance selection of the A- or B-object and in the long run its selection approximates chance performance. This trial is not represented in the figure. The addition of three more objects on the second trial causes *Ss* to err, *i.e.* choose an object other than A, on 66.7% of the problems. On the third trial *Ss* are again given the two objects and make 25.7% errors.

The right-hand side of the figure shows *S*'s performance, on an average

[3] R. S. Massar and R. T. Davis, The formation of a temporal sequence learning set, *J. comp. physiol. Psychol.*, 53, 1959, 225-227.

problem, during the second 14-day period of the experiment. Between the first and second halves of the experiment all *S*s show improvement in performance ($p < 0.01$ using the sign-test). The *S*s in the control group improve most strikingly on the trials containing five stimuli, and on these trials their performance is significantly better than that of the irradiated *S*s.[4]

The conspicuous difference in level of performance between trials employing two and five objects led the writer to suspect that two different processes were operating. The rank-order coefficient of correlation between *S*s' total scores on trials with two and five objects was, however, 0.967.

The data were reanalyzed and sorted according to six categories of response to determine if *S*s in the two groups made the same kind of errors, *i.e.* were equally disposed to select the B-object and one of the N-objects. Two categories, designated '+' and '−,' were determined by the reward or nonreward respectively of *S*s' preferences on Trial 1. Both of these categories were divided into three subcategories, each of which pertained to the relative frequency that *S* selected a particular object from the five objects present on Trial 2 or subsequent critical trials. These cubcategories included responses on Trial 2 to: (a) the object designated correct (A) by *E* on Trial 1; (b) the object *E* designated incorrect (B) on Trial 1; and (c) one of the three objects designated neutral (N). The results of this analysis indicated that *S*s selected +A significantly more often than −A or −N and responded significantly more frequently in these categories than in the three remaining categories, +N, −B or +B. These relationships held for Trial 2 and all subsequent trials and did not change with practice. Further analysis of the incorrect responses indicated that experimental and control groups make different kinds of incorrect responses and that the size of this difference depended on the object selected on the previous trial. Fig. 2 shows the percentage of the incorrect responses that were made to the B-object in preference to the alternative incorrect objects and relates the selection of B-objects to the selection of the A- or B-object on Trial 1 (righthand side of figure) and on all odd numbered trials, (left-hand side of figure). Since this comparison does not involve the A-object, chance is 25% rather than 20%. The computation of points in this figure involved a ratio between the number of times the B-object was selected under each of the conditions represented in the abscissa and the number of errors made under comparable conditions.

---

[4] The $p < 0.05$ by the Mann-Whitney rank-test. See S. Siegel, *Nonparametric Statistics,* 1956, 120.

The figure indicates that *S*s choose the *B*-object less frequently than one of the alternative incorrect N objects, particularly after selecting the B-object on an odd-numbered trial. If *S* has selected the A-object on an odd-numbered trial, it is more likely to select the B-object than one of the N-

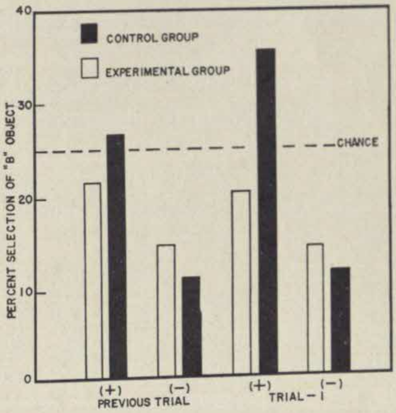


FIG. 2. FREQUENCY OF ERRORS ON EVEN-NUMBERED TRIALS THAT ARE THE B-OBJECT RATHER THAN ONE OF THE ALTERNATIVE NEUTRAL OBJECTS Prior success and failure and experimental treatment are indicated.

objects and this tendency is especially marked among animals in the control group following a single trial.

Summary. During the last half of the experiment irradiated monkeys performed significantly less well than nonirradiated animals on trials which required *S* to choose one out of five objects. There was no significant difference in *S*s' scores on trials on which they discriminated between two stimuli. If the A-objects is not chosen on trials with five objects, one of the alternative objects is chosen in preference to the B-object if the prior trial was an error. When *S*s in the control group err, they are more likely to select the B-object than one of the alternative objects.

VOCAL CUES TO THE IDENTIFICATION OF LANGUAGE

ALAN COHEN, Northwestern University, and JOHN A. STARKWEATHER, University of California School of Medicine

For somewhat more than a decade, a small series of studies has sought to answer an interesting question about ethnocentric prejudice. This question is whether the prejudiced person, because of a greater need to identify those of different racial identity, becomes more sensitive to subtle cues. One of the first of these studies stated the possibility this way. "The question of racial identity is of small importance to the person free from prejudice, yet it is of considerable importance to the bigot, and for this reason, the bigot apparently learns to observe and interpret both facial features and expressive behavior so that he can more swiftly spot his 'enemy'."[1] The studies referred to used photographs in order to attack the question of the interpretation of facial features. They have produced conflicting results about any special ability of the prejudiced person to better use such cues.[2] In general, the prejudiced subjects placed a minority group label on more photographs and tended to score more accurately only if the majority of the pictures were of minority group members.

Our attempt was to attack this same question but with regard to recognition of a specific type of *expressive* behavior. There are situations where one has clearly heard the words of someone conversing in a foreign language. The words are obvious indications to most listeners of a national difference. Could the listener make this identification, however, if he were not able to make out the words but only heard some vocal characteristics of the conversation? For example, if he were in an adjoining room and heard the voices muffled through the wall?

This, of course, is not a new area of interest. Recall the central part it played in one of Edgar Allen Poe's mysteries, *The Murders in the Rue*

* Received for publication March 2, 1959. The research was done at the University of California, Berkeley.

[1] G. W. Allport and B. M. Kramer, Some roots of prejudice, *J. Psychol.*, 22, 1946, 17.

[2] See, for example, F. H. Lund and W. C. Berg, The identifiability of national characteristics, *J. soc. Psychol.*, 24, 1946, 77-83; L. J. Carter, The identification of racial membership, *J. abnorm. soc. Psychol.*, 43, 1948, 279-286; Gardner Lindzey and Saul Rogolsky, Prejudice and identification of a minority group membership, *ibid.*, 45, 1950, 37-53; D. N. Elliot and D. H. Wittenberg, Accuracy of identification of Jewish and non-Jewish photographs, *ibid.*, 51, 1955, 339-343.

Morgue. In this story, a number of witnesses testified about sounds they heard from outside the room where the murders were committed. These witnesses, who were of different European nationalities, all described one voice as that of a foreigner. Many were very sure it was a particular language from the intonation, but they all described a language which they did not themselves understand. They were, however, all in error, since the voice was actually that of an orangutan.

To return to the original question, there may be *some* to whom the identification from voice is particularly important and who have found ways to identify language on the basis of vocal cues alone.

Method. Short recorded samples were prepared of 14 different speakers, talking in their native tongues and representing five languages (Chinese, English, German, Italian, Spanish). The voice-samples were played in scrambled order to 60 undergraduate students (*S*s) who worked in small groups. They first chose between 10

TABLE I

PERCENTAGE OF CORRECT IDENTIFICATION OF LANGUAGE FROM FILTERED AND CLEAR RECORDINGS

*S*s	Filtered		Clear	
	Non-English	English	Non-English	English
High Ethnocentric ($N=15$)	23.6	73.4	67.3	97.7
Total ($N=60$)	23.6	66.0	69.2	97.8

alternatives of favorable or unfavorable adjectives to describe each sample, and then during a second playing guessed the language being spoken. The first two presentations were filtered through a 600 ~ low-pass filter which removed the *verbal* speech *content* and left only *vocal* information in the recordings. The filter and its effects on intelligibility have been previously described.[3] During a third and fourth presentation of the voices, the *S*s again chose between alternatives of favorable and unfavorable adjectives and again guessed the language being spoken. During these presentations the voices were heard normally, clear and unfiltered. Two measures were used as estimates of the *S*s' racial prejudice: 20 items from the California Ethnocentrism Scale and the difference in number of unfavorable adjectives applied to non-English versus English voice samples.[4]

Results and discussion. Table I indicates the proportion of correct language-identifications for the two conditions of filtered and clear recordings, and for the average of 11 non-English voices and three English voices

[3] J. A. Starkweather, The communication-value of content-free speech, this JOURNAL, 69, 1956, 121-123; Content-free speech as a source of information about the speaker, *J. abnorm. soc. Psychol.*, 52, 1956, 394-402.

[4] T. W. Adorno, Else Frankel-Brunswik, D. J. Levinson, and R. N. Sanford, *The Authoritarian Personality*, 1950, 102-150.

under the two conditions. English was identified significantly more often than non-English samples, both when verbal content was removed by filtering and when all the normal cues were present ($p < 0.01$). Perhaps more surprising than the ability to make identifications when using only vocal cues is how 4 out of the 60 *S*s could miss in the identification of English in what were quite clear recordings. The over-all rate of guessing English was 26%.

Twenty percentage would be the chance guessing level if the *S*s had been given the five languages which were present and asked to choose among them. Our judges were engaged in free guessing, however, without this information, and the true chance-level was therefore lower than this to an unknown degree. The results indicate, however, that even the vocal cues alone allowed some increase over chance expectation, though it

TABLE II

PERCENTAGE OF UNFAVORABLE ADJECTIVES USED TO DESCRIBE FILTERED AND CLEAR VOICES

*S*s	Filtered		Clear	
	Non-English	English	Non-English	English
High Ethnocentric ($N=15$)	47.4	30.4	38.0	28.6
Total ($N=60$)	42.1	28.1	33.7	24.3

was not impressive for non-English. Because of the large number of judgments involved, even the identification of non-English voices was statistically above the chance-level ($p < 0.05$).

With regard to our original question about prejudice, the results seem quite negative. The *S*s who obtained high scores on the Ethnocentrism Scale (the upper 25% of our group) did not differ from these over-all in their ability to identify. The same is true for the *S*s who showed the greatest differential in the number of unfavorable adjectives which they attributed to non-English versus English voices. There is some tendency (not significant statistically) for High E *S*s to identify filtered English more often. The two measures with which we hoped to estimate racial prejudice were not related to each other, having a correlation of 0.06.

The proportion of unfavorable adjectives checked by the *S*s is shown in Table II for the two conditions. Here again there are differences in both filtered and clear conditions between non-English and English voices, with more unfavorable adjectives checked about the non-English voice-samples. Again the *S*s with relatively high ethnocentrism scores produced essentially the same data as other judges.

We were not successful in showing any special sensitivity of the prejudiced person with respect to this kind of expressive behavior. There are many possibilities, of course, as explanation for such a failure without necessarily abandoning the original hypothesis. It may be, for example, that we have chosen an expressive behavior that is seldom found as a natural occurrence. The data indicate, however, that the vocal portion of speech carries information which allows some recognition of language. For English-speaking listeners it allows a high degree of recognition of English versus other languages.

THE EFFECT OF DANGER UPON THE EXPERIENCE OF TIME

By Jonas Langer, Seymour Wapner, and Heinz Werner,
Clark University

Studies on the interrelationship of space and time have demonstrated that when the time-interval between successive flashes of lights is kept constant while the physical distance between them is varied, the observer's experience of elapsed time does not remain constant but varies with the physical distance.[1] A further question may be asked as to whether such space-time interrelationship also obtains if the psychological distance is changed through variations other than those of physical distance.

Werner and Wapner developed a method of obtaining changes in psychological distance with physical distance held constant. They found that psychological distance changed under conditions of danger; *e.g.* when *S* walked toward and stopped short of a precipitous edge, he overestimated the distance he travelled, or stated another way, the edge appeared closer. The introduction of danger affected psychological distance even though physical distance was not changed.[2]

The purpose of this study was to determine whether the space-time interrelationship obtains when psychological space alone is changed. The presence of danger, known to affect psychological distance, was introduced as the experimental condition. It was thought that changes in psychological distance, under conditions of danger, would be paralleled by changes in psychological time. More specifically, it was expected that since distance traversed is overestimated under conditions of danger, time elapsed would be overestimated under danger.

Procedure. S's task was to judge a specific time-interval while being moved toward and away from a precipitous edge. The apparatus (Fig. 1) consisted of a platform on wheels with handrails. It was steered by *E* from behind, and driven by an electric

* Received for publication September 26, 1959. This investigation was supported in part by PHS Research Grant M-348 from the National Institute of Mental Health of the U. S. Public Health Service.

[1] John Cohen, C. E. M. Hansel, and J. D. Sylvester; A new phenomenon in time judgment, *Nature,* 172, 1953, 901-903, Interdependence in judgments of space, time, and movement; *Acta psychol.,* 11, 1955, 360-372. For the reverse effect of time on distance, see Harry Helson and S. M. King, The *tau* effect. An example of psychological relativity, *J. exp. Psychol.,* 14, 1931, 202-217; and P. E. Comalli, The effect of time on distance perception. Unpublished Master's thesis, Clark University, Worcester, Massachusetts, 1951.

[2] Heinz Werner and Seymour Wapner; Changes in psychological distance under conditions of danger, *J. Pers.,* 24, 1955, 153-167.

motor at a constant speed of two miles per hour varying slightly with the weight of *S*.

S stood on the platform with his hands on the rails. A button on the right handrail enabled *S* to activate the motor which drove the platform; *S*'s button did not operate unless *E*'s master switch was also turned on. A Standard Electric Timer was wired into the circuit which measured the time the button was depressed.

The experiment was performed along a corridor at the end of which was a stairwell. *S* stood on the platform with the toes of his shoes at its forward edge. He wore a blindfold which he had to raise before each trial. *S*'s task was to indicate an

FIG. 1. SKETCH OF APPARATUS

interval of 5 sec. by pressing and releasing the button. Depressing the button started the movement of the platform and simultaneously the timer; releasing it simultaneously stopped both.

Judgments of the time-interval were made under two conditions, 'Danger' and 'No danger.' 'Danger' involved moving blindfolded *toward* the stairwell. There were two trials involving two degrees of danger; one started 15 ft. from the edge (more dangerous), the other 20 ft. from the edge (less dangerous). 'No danger' involved judgments of the time-interval while *S* was moving, blindfolded but *away* from the stairwell. Two trials, one starting at a distance of 15 ft. and the other 20 ft. from non-dangerous markers, were performed. Thus, a total of four measures was obtained for every *S*. The sequence of trials was varied in accordance with a 4 $\times$ 4 Latin Square. Sixteen *Ss*, eight men and eight women, were tested.

Instructions. Before the start of the experiments, *S* was instructed as follows:

Your task throughout will be to judge when 5 sec. are up. It is important that you do not use any method, such as counting, to judge the passage of time. [Before each trial, *S* was specifically instructed:] Raise the blinders, please. From the time I say 'Go' I wish you to press the button which will start you moving [in the 'Danger' condition] toward the edge of the stairs; [in the 'No danger' condition] toward the first or second marker [depending on the trial in the sequence]. When you think 5 sec. are up, lift your finger off the button and you will stop. Remember, do not use any method such as counting. Lower the blinders, please. 'Ready—Go.'

Results. The results are in keeping with the hypothesis that as danger is increased, the actual time (as measured by a clock) judged by *S* to be equivalent to a pre-specified time-interval, decreases. Table I shows that under 'No danger' an over-all mean of 4.11 sec. was judged equivalent to 5 sec., whereas under 'Danger' the mean was 3.52. The difference, 0.59 sec., is significant at the 1% level.

A comparison of the two 'Danger' trials indicates, in accordance with expectation, that as danger increases the overestimation of time increases.

TABLE I

EFFECT OF DANGER ON THE ESTIMATION OF A 5-SEC. INTERVAL

Sources of variance	*df.*	Mean square	*F*	*P*
Between *S*s:	15	6.71		
sequence	3	5.82	0.84	>0.05
residual	12	6.93		
Within *S*s:	48	0.74		
combined treatments	3	2.23	3.38	<0.05
'Danger'-'No danger' (D)	1	5.57	8.45	<0.01
distance position (P)	1	0.02	0.03	>0.05
D×P	1	1.09	1.65	>0.05
order	3	0.98	1.49	>0.05
square uniqueness	6	0.34	0.52	>0.05
residual	36	0.66		
Total	63			

	Mean elapsed time (sec.)		
Distance	'No danger'	'Danger'	Diff.
15 ft.	4.22	3.37	0.85
20 ft.	4.00	3.67	0.33
overall mean	4.11	3.52	0.59

The mean time-judgment, when started 20 ft. from the edge of the stair-well was 3.67 sec.; when started 15 ft. from the edge it was 3.37 sec. The difference (0.30 sec.), though not significant, is in the expected direction.

Conclusion. The major results of this study are that time is overestimated during danger and that the overestimation tends to increase as danger increases. This finding confirms the hypothesis that changes in psychological distance are paralleled by changes in psychological time even when physical distance is held constant. They also confirm the findings of Werner and Wapner that the introduction of affective conditions, such as danger, changes the experience of distance.[3] When the experience of distance is

[3] For a fuller theoretical discussion, the reader is referred to Werner and Wapner, Changes in psychological distance under conditions of danger, *J. Pers.*, 24, 1955, 159-167.

changed under conditions of danger, with physical distance held constar the concomitant experience of time is also changed. While *S* overestimat the distance he has moved under conditions of danger, he also overes mates the time that has elapsed. As danger is increased, *S* experienc movement through an increasingly longer distance; concomitantly, he e periences that an increasingly longer interval of time has elapsed, at ea point, than actually, as measured by a clock, has occurred at that point.

In general, this means that as danger increases, space and time stretc or perception of a given physical distance or interval of time shrinks.

APPARENT VISUAL SIZE AS A FUNCTION OF DISTANCE FOR MENTALLY DEFICIENT SUBJECTS

By HERSCHEL LEIBOWITZ, University of Wisconsin

A number of experiments have demonstrated less size-constancy in children than in adults, especially when the distance of the test-object is increased.[1] This difference may be attributed either to the greater experience of adults with such objects or to their greater intellectual development. The purpose of the present study is to examine the relative merits of these alternative explanations by comparing the performance of institutionalized mental defectives with that of normal *S*s of approximately the same chronological age.

Subjects. Twenty 'familial' defectives from the resident population of the Southern Wisconsin Colony and Training School comprised the mentally deficient group. None had any history of brain damage, epilepsy, eye-defect, or other organic pathology. In a preliminary experiment, in which an attempt was made to test *S*s of the lowest possible mental age, it was discovered that none of the *S*s with mental ages below 7 yr. could understand or follow the directions. The mean *MA* of the group successfully tested was 8.7 yr., with a range of 7–12 yr. Their *CA*s ranged from 14–27 yr. with a mean of 21.3 yr.

The 10 normal *S*s were undergraduate students at the University of Wisconsin who were transported to the Colony and tested under the same conditions as the mental defectives. Their age-range was 19–26 yr., with a mean of 21.2 yr.

Method. The apparatus and procedure were designed to be as simple as possible. *S* sat at the junction of two well-illuminated corridors located at right angles to each other, and signaled which of two vertical sticks, one visible on the floor of each corridor, looked the longer. The test-objects were wooden dowels cut from 1-in. stock to lengths of 2, 4, 8, 16, and 24 in. and mounted individually in square bases $4 \times 4 \times 1$ in. Both the dowels and the bases were painted flat black. The test-objects were placed, one at a time, at viewing distances of 10, 20, 40, 80, and 120 ft., respectively. Thus, each subtended the same visual angle (0.96°). The comparison-stimuli, which were constructed in the same manner as the test-objects, were also presented individually, but at a fixed viewing distance of 20 ft. The

* Received for publication November 13, 1959. This work was supported in part by Grant M1090 from the National Institute of Mental Health. The cooperation of the staff of the Southern Wisconsin Colony and Training School, Union Grove, Wisconsin, and the aid of Mr. Philip McGuire and Mr. Frederick Bussey are gratefully acknowledged.

[1] Franz Beyrl, Ueber die Grössenauffassung bei Kindern, *Z. Psychol.*, 100, 1926, 344-371; Wilhelm Burzlaff, Methodologische Beiträge zum Problem der Farbenkonstanz, *ibid.*, 119, 1931, 177-235; J. Piaget, *The Psychology of Intelligence*, 1950, 53-64; H. P. Zeigler and Herschel Leibowitz, Apparent visual size as a function of distance for children and adults, this JOURNAL, 70, 1957, 106-109.

comparison-series ranged in length from 1 to 36 in. in ½-in. steps to 6 in., and 1-in. steps for the longer stimuli.

The procedure consisted first of asking *S* if he would like to help judge t sizes of sticks and explaining that the task consisted simply of pointing to the sti which "looks taller." The comparison-stimuli presented initially were so differe in size from the test-object that a discrimination could be elicited easily. It w considered important to shift the initial discriminations from a very large to very small comparison, or vice versa, to be certain that the instructions were und stood. A number of such checks were made during the course of testing the feeb minded *Ss*. Otherwise, the method of limits was employed with an ascending and

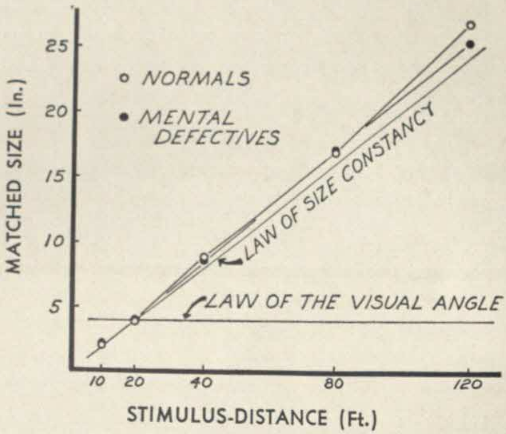


FIG. 1. MEAN MATCHED SIZE AS A FUNCTION OF THE DISTANCE OF TEST-OBJE FOR GROUPS OF MENTAL DEFECTIVES AND OF NORMALS.

The test-object subtended the same angle at all distances of observation.

descending series for each test-object. The order of presentation of the test-objec was determined by a 5 × 5 Latin Square design replicated four times for the me tal defectives and twice for the normals. No time-limit was placed on *S*. All sessio were completed within 15 min.

Results. The mean matched size of the comparison-stimulus as a functio of the distance of the test-objects is plotted in Fig. 1. On this plot, the hor zontal line represents a prediction in terms of the 'law' of the visual angl Since the test-objects all subtended the same visual angle, agreement wit this prediction would demand the same matched size independent of view ing distance. The 'law' of size-constancy is represented by the diagonal lin which is determined by plotting the actual sizes of the test-objects. Th data indicate that both groups judged the sizes of the test-objects correctl at all distances despite the fact that the retinal images they produced di not vary.

Discussion. The function relating apparent size and distance for the normal *S*s lies very close to the prediction in terms of size-constancy, as do similar functions obtained with normal adults utilizing free binocular vision over a wide range of viewing distances.[2] The feeble-minded group, despite mental retardation, demonstrate a similar function which is essentially indistinguishable from that produced by the normals. Both groups match correctly the sizes of the test-objects at the various viewing distances.[3] These data imply that the results of previous experiments with normal children are to be attributed to experiential rather than to intellectual factors. Stated differently, whatever factors are responsible for mental retardation do not interfere with the normal development of size-constancy. Jenkin and Morse, in an experiment similar in many respects to the present study, also report no difference between groups differing in *MA* but of approximately the same *CA*.[4]

The same feeble-minded *S*s tested here had previously participated in an experiment on shape-matching in which large differences were found in the extent to which mentally deficient *S*s exhibit shape-constancy as compared with normals.[5] The higher the intellectual level of the group, the less was their tendency to respond in terms of shape-constancy. The absence of such differences for size-matching emphasizes the point of view that the different constancies are mediated by different mechanisms.

Summary. The relation between perceived size and distance was determined for a group of familial mental defectives and a group of normal *S*s of the same *CA*. There was no difference in the performance of the two groups. The results support the view that the development of size-constancy in normal children is a result of experiential rather than intellectual processes.

---

[2] A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37; A. S. Gilinsky, The effect of attitude upon the perception of size, this JOURNAL, 68, 1955, 173-192; Herschel Leibowitz, P. Chinetti, and J. Sidowski, *Science*, 123, 1956, 668-669; Zeigler and Leibowitz, *op. cit.*, 106-109.

[3] Actually, both groups matched the test-object with a comparison-stimulus slightly larger than would be predicted in terms of size-constancy.This is an example of the phenomenon of 'overconstancy' discussed in greater detail by Holway and Boring, *op. cit.*, 27; J. J. Gibson, Motion picture testing and research, Report No. 7, Army-Air Force Aviation and Research Reports, 1947, 201-212; W. M. Smith, A methodological study of size distance perception, *J. Psychol.*, 35, 1953, 143-153; Gilinsky, *op. cit.*, 186; Noël Jenkin, Effects of varied distance on short-range size judgments, *J. exp. Psychol.*, 54, 1957, 327-331.

[4] Noël Jenkin and Sally M. Feallock, Developmental and intellectual processes and size-distance judgment, this JOURNAL, 73, 1960, 168-273.

[5] Herschel Leibowitz, I. Waskow, N. Loeffler, and F. Glaser, Intelligence level as a variable in the perception of shape, *Quart. J. exp. Psychol.*, 11, 1959, 108-112.

THE SUBJECTIVE EDGE

By B. M. SPRINGBETT, University of Manitoba

Marshall and Talbot offer a general theory of contour-perception which fairly covers a wide range of perceptual phenomena.[1] One set of deductions made by them concerns the nature of perception of a hair-line contour as compared with an edge—a contour formed by adjacent areas differing in intensity, or hue, or both.[2] The predicted phenomena do not, however, appear to have been demonstrated empirically. According to the theory, the steepest gradient of neural excitation, which defines or locates the position of the perceived contour, should lie toward the dark side of the edge, *i.e.* the perceived, or subjective, edge should lie 'into the dark' from the true or geometrical edge, while the preceived hair-line contour should be coincidental with the geometrical position of the contour. Furthermore, because of the essentially sigmoidal distribution of excitation postulated for the edge-contour and the gaussian distribution of the hair-line contour, fluctuations in the locus of the subjective edge will occur while the perceived hair-line contour will remain stable or, at least, will fluctuate less markedly.

The following experiment is designed to test these deductions.

Method. The apparatus consists of a small metal cabinet in which two 7-w. white lamps are mounted and faced with plexiglass to diffuse the light. Each lamp independently illuminates a slide from behind. Slides fit into apertures cut in the face of the cabinet.

The two slides are vertically separated by 0.5 in. The lower slide is in a fixed position, the upper moves horizontally. The position of the upper slide is adjustable by turning a conveniently situated knob. A gear-system permits fine adjustments, *i.e.* 1° of turn moves the slide 0.014 in. A 4-in. pointer at *E*'s end of the knob-spindle permits accurate reading of the position of the upper slide.

Three pairs of slides are used. One pair has a vertical hair-line contour, while the remaining two pairs have vertical contours (edges) formed by adjacent areas, one black-white, the other blue-yellow. The positions of the areas in the upper and

---

* Received for publication February 28, 1960. This work was supported by grant APBT-40 from the Associate Committee on Applied Psychology of the National Research Council of Canada.

[1] W. H. Marshall and S. A. Talbot, Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity, in Heinrich Klüver (ed.) *Biological Symposia,* VII, 1942, 117-164; C. E. Osgood, *Method and Theory in Experimental Psychology,* 1953, 229-260.

[2] Marshall and Talbot, *op. cit.,* 138-141.

lower slides of these pairs are reversed, *e.g.* the lower black-white slide has the black on the right, the upper one on the left. Thus, when the upper slide is set to the right of a true-zero alignment and *S* is asked to align the upper and lower contours, under-estimation is predicted; when the setting is to the left, over-estimation is predicted. Errors on the hair-line are theoretically required to occur randomly with respect to over- and under-estimation. Phenomenal fluctuation of contours will be reflected in the variance of error-scores.

With *S* seated, the test-object is presented at eye-level, at a distance of 24 in. A black cardboard 'tunnel' provides a uniform surround within which the slides are

TABLE I

RATIOS OF CORRECT TO INCORRECT PREDICTIONS

Predicted	Obtained					
	black-white		blue-yellow		hair-line	
	+	−	+	−	+	−
Over-estimation (+)	171	29	191	9	41	59
Under-estimation (−)	27	173	7	193	29	71
Chi square (1 *df.*)	204.68		334.42		2.66	

viewed. After a period of 5 min. pre-adaptation in a darkened room, *S* makes 20 judgments on each of the edge-contours and 10 on the hair-line contour. The *S*s are 20 university students.

Results. In studying the locus of the subjective edge, two bases of analysis are used: (a) the population of judgments, $N = 400$ for each edge-contour, 200 for the hair-line contour; (b) the population of *S*s, $N = 20$ for each contour. Table I shows the ratio of correct to incorrect predictions. With settings made to the left of zero, over-estimation (+) is predicted; under-estimation (−) is predicted with settings to the right. There seems to be little doubt that systematic errors occur in the predicted directions with the edge-contours. There is no evidence of systematic error in the hair-line, as predicted.

The judgments of each *S* were summed algebraically and averaged to yield a measure of constant error. The *SD*s reported relate to inter-*S* variability. Table II shows the mean (constant) error for the group on each contour, together with *SD*s and SE_Ms.

Calculations made from the table show that the distances of the means from zero, *in terms of* SE_M, are 0.88, 12.00, and 13.33, respectively, for the hair-line, black-white, and blue-yellow contours, *i.e.* the hair-line does not differ significantly from zero, while the edge-contours do. The difference between the hair-line and black-white means is significant ($t = 9.939$); for hair-line vs. blue-yellow, $t = 11.47$; and the blue-yellow mean is significantly

greater than the black-white ($t = 3.346$). The last difference between edge-contours cannot be explained in terms of the present data, but a suggestion may be made about it.

In each set of data (black-white, blue-yellow) the 10 judgments from the lefthand settings may be averaged for each *S*, and, similarly, the 10 from righthand settings. This provides a rather crude basis for estimating a split-half reliability coefficient and the resulting *r*s are, for black-white, 0.735; for blue-yellow 0.756. Next, the average black-white score for each *S* is correlated with the average score on the blue-yellow ($r = 0.196$). Each reliability-coefficient is significantly different from that of the inter-correlation: the value of t for black-white vs. inter-correlation is 2.167; the corresponding value of t for blue-yellow is 2.281. Each difference is significant at the 5% level. The suggestion is that the black-white and blue-yellow judgments do not belong to the same population, *i.e.* different sources of variance are operating in each.

TABLE II

LOCUS AND VARIABILITY OF THREE DIFFERENT CONTOURS

Contour	Constant error (in.)	SD	SE_m
Hair-line	0.0016	0.0088	0.0018
Black-white	0.036	0.013	0.003
Blue-yellow	0.046	0.016	0.0036

Fluctuations in the subjective edge as compared with the hair-line contour are studied over a population of judgments; $N = 400$ for black-white and blue-yellow; $N = 200$ for the hair-line. The *SD*s for these respective sets of data are 0.029 in., 0.027 in., and 0.019 in. By Guilford's formula,[3] the *SD* for the black-white data is found to be significantly greater than that of the hair-line ($t = 5.248$). In a similar comparison for blue-yellow, $t = 5.265$. The results are in agreement with the postulated greater variability of the edge as compared with the hair-line.

While the Marshall-Talbot deductions are in general supported by these data, a comparison of the results with black-white with those with blue-yellow raises questions which can find answers only through additional experiments. Phenomena associated with the subjective edge should be investigated with intensity-ratio as the experimental variable. Similarly, with intensity-ratio constant, differences in hue should be explored in relation to their ability to produce a subjective edge.

[3] J. P. Guilford, *Fundamental Statistics in Psychology and Education*, 2nd ed., 1950, 197.

APPARENT WEIGHT AS A FUNCTION OF HUE

By M. Carr Payne, Jr., Georgia Institute of Technology

In an earlier study, we found that different colored objects of the same size differ in apparent weight.[1] These differences were related to physical brightness. No relation was found between apparent weight and preferences for the colors. The present study was undertaken to test three hypotheses: (1) that objects of the same size but of different hues differ in apparent weight; (2) that the apparent weight of an achromatically colored object differs from that of a chromatic object of the same brightness; and (3) that there is no significant relation between apparent weight and preference for hue.

Method. Six cubes, 2 in. on a side, used in the earlier study, were covered with Munsell papers of equal brightness and saturation but of different hues. Red, yellow, green, blue, purple, and gray papers were used.

Cubes were presented to *S* by the method of single stimuli. Order of presentation was randomized with the restriction that each of the six hues was presented before any of them was repeated. Each cube was presented 15 times to every *S*.

S saw each cube in the center of a 6-in. aperture. The cube was presented on a white base against an evenly illuminated white background. The aperture was cut from white cardboard and was surrounded by white sheeting in such a way as to reduce possible assimilation effects.[2] The surround of the aperture was diffusely illuminated to a level slightly below that of the cube.

Twenty-five college students (men) served as the *S*s. Each, serving individually, was seated 10 ft. 2 in. from the aperture and instructed to call out a number from 1 through 5 as each cube was presented according to its apparent weight. *S* was given a sheet with a scale and descriptive adjectives to help him assign the numbers. As soon as *S* made his judgment the aperture was closed. When it was reopened, *S* saw another cube.

To insure *S*'s understanding of the instructions a white cube was first presented him and he was asked to judge its apparent weight. The first trial with every cube of different hue was similarly treated as practice.

After 15 test-trials with each cube had been completed, *S* was shown a card containing a 2-in. square of each of the hues. He was instructed to rank these hues in order of their preference.

---

* Received for publication January 28, 1960.

[1] M. C. Payne, Jr., Apparent weight as a function of color, this Journal, 71, 1958, 724-730. Cf. also C. J. Warden and E. L. Flynn, The effect of color on apparent size and weight, this Journal, 37, 1926, 398-401.

[2] P. V. Marchetti, Time-errors in judgments of visual extents, *J. exp. Psychol.*, 30, 1942, 257-261; D. C. McClelland, Factors influencing the time error in judgments of visual extent, *ibid.*, 33, 1943, 81-95.

Results. After testing for homogeneity of variance by Hartley's test,[3] an analysis of variance was computed. Results of this analysis are shown in Table I. Interactions for Hue × *S* and Trial × *S* were significant ($P < 0.01$) when tested against the second order interaction as an error-term. Hues and *S*s were significant ($\rho < 0.01$) as main effects when tested against the H × *S* interaction.

To evaluate the main effects of Hue, critical ratios were computed between mean apparent weights for hue. Yellow, green, and gray did not differ significantly from one another. Blue, red, and purple did not differ significantly from one another. The apparent weight of yellow, green, and gray did, however, differ significantly from that of blue, red, and purple.

TABLE I

ANALYSIS OF VARIANCE

Source of variation	*df.*	Mean square	*F*
Hues (H)	5	38.31	3.70*
Subjects (*S*)	24	27.22	2.63*
Trials (T)	14	.52	1.79
H×*S*	120	10.34	36.93*
H×T	70	.29	1.04
T×*S*	336	.32	1.14†
H×T×*S*	1680	.28	

* $\rho<0.01$; † $\rho<0.05$.

To determine if there is a significant relation between apparent weight and preference for hue, a rank order correlation coefficient was computed between the ranks of preference and those of apparent weight. This correlation was not significant ($\rho = 0.36$).

Discussion. The present study indicates that the hue of an object does influence the perceived weight of an object. The extent of this influence does not appear, however, to be as great as that of brightness (reflectance) which was shown by the earlier study to correlate very greatly ($\rho = 0.94$) with apparent weight. This conclusion is further supported by the lack of a significant difference between the apparent weight of the gray cube and that of a green or yellow cube of equal brightness. Apparently, brightness is the major cue for apparent weight.

[3] H. M. Walker and Joseph Lev, *Statistical Inference,* 1953, 192.

APPARATUS

A MONOPOLAR METHOD OF MEASURING PALMAR CONDUCTANCE

By Robert B. Malmo and John F. Davis, McGill University

The apparatus and method described in this paper have been successfully used in a number of studies conducted in our laboratory, and we wish to make them available to our co-workers.[1] The monopolar method (which makes use of an indifferent or reference electrode and an active electrode) was selected because we wished to study the variations of skin-conductance at a single site.

Fig. 1 presents the circuit diagram for the apparatus. It is convenient to work with resistance units during recording, converting to conductance-units later on. A subject's (*S*'s) resistance is the unknown factor in the Wheatstone bridge that is approximately balanced by means of a General Radio Type 1432 Decade Resistance Unit, the unbalanced portion being recorded by the direct coupled (*DC*) amplifier and recording galvanometer. Three decades, in units of 1,000, 10,000, and 100,000 Ω will serve all requirements for recording from *S* (a separate millivoltmeter for checking the electrodes will be described later).

For purposes of recording the palmar conductance (*PC*), the direct coupled (*DC*) amplifier should have maximal available gain of 200 (al-

* Support for the development of this method has come from the following sources: Medical Research and Development Division, Office of the Surgeon General, Department of the U.S. Army: Contract Number DA-49-007-MD-626; National Institute of Mental Health, National Institutes of Health, U.S. Public Health Service: Grant Number M-1475; the Department of National Health and Welfare (Canada); Defence Research Board, Department of National Defence, Canada: Grant Number 19425-04; and the National Research Council of Canada: Grant Number A.P. 29.

Grateful acknowledgment is made to Mr. W. Mundl for technical assistance.

[1] S. M. Feldman, Differential effect of shock as a function of intensity and cue factors in maze learning. Unpublished Doctoral dissertation, McGill Univer., 1958; R. B. Malmo, Measurement of drive: An unsolved problem in psychology, in M. R. Jones (ed.), *Nebraska Symposium on Motivation, 1958.* 1958, 229-265; Activation: A neuropsychological dimension, *Psychol. Rev.,* 66, 1959, 367-386; M. M. Schnore, Individual patterns of physiological activity as a function of task differences and degree of arousal, *J. exp. Psychol.,* 58, 1959, 117-128; R. G. Stennett, The relationship of performance level to level of arousal, *J. exp. Psychol.,* 54, 1957, 54-61; R. G. Stennett, The relationship of alpha amplitude to the level of palmar conductance, *EEG Clin. Neurophysiol.,* 9, 1957, 131-138.

though gain used at the sensitivity that we generally employ is only 33), and its drift should be no more than 2 mv. per hr. Since such amplifiers are not readily available from commercial sources (whose amplifiers have too much or too little gain), one must either be specially constructed or else converted from one of the models available. At present, we use an Epsco-type 8105 *DC* amplifier in which we have reduced the gain, by a factor of 10 or 15 to 1, by eliminating one of the stages.

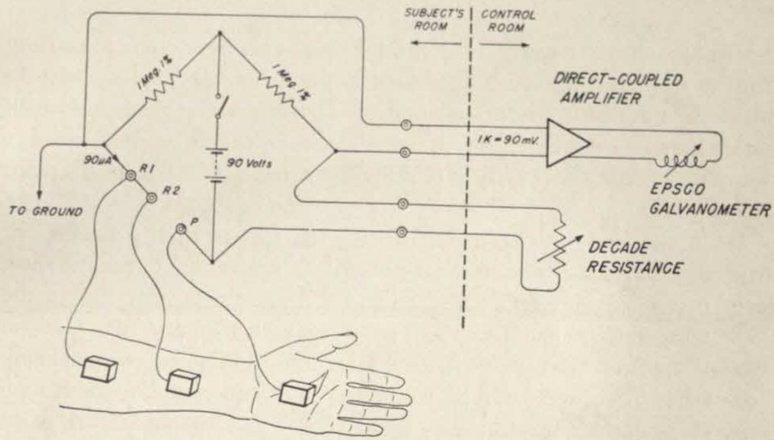


FIG. 1. CIRCUIT DIAGRAM FOR SKIN-RESISTANCE RECORDING.
R1 and R2 designate the two reference electrodes. See text for explanation.

If the current through *S* is too small, then the voltage drop through the skin resistance might be of the same order of magnitude as the skin voltage (*i.e.* the more or less steady biological potential which is always present across the skin). This makes it difficult to differentiate between skin voltage (Tarchanov) effects and skin-resistance (Fere) effects, and we wished to study the latter effects exclusively. On the other hand, if the current is too large, there is always the danger of polarization of the electrodes. We selected 90 μA. as a compromise which avoids both difficulties. Since our bridge circuit is symmetrical, the current is the same in each leg.

In arranging the apparatus, if two adjoining rooms are available, it is preferable (less distracting for *S*) to install the Decade Resistance Unit, *DC* amplifier, and chart drive in a control room away from *S*.

Connecting PC reference leads to ground. With multiple-channel recording, especially when electromyograms (*EMGs*) are included, it is desirable

to ground *S* to prevent 60-cycle interference from various sources. In the absence of this ground, it is often impossible to secure *EMG*s at high gain that are free from artifacts, and sometimes the 60-cycle artifact is present even in the *PC*-tracing. Fig. 1 shows the reference leads going to ground (there are two ground electrodes connected in parallel).

It should be noted that grounding the body at more than one point is decidedly unwise. With more than one ground, and with the *PC*-bridge grounded at some point, the current that is being passed through *S* will be

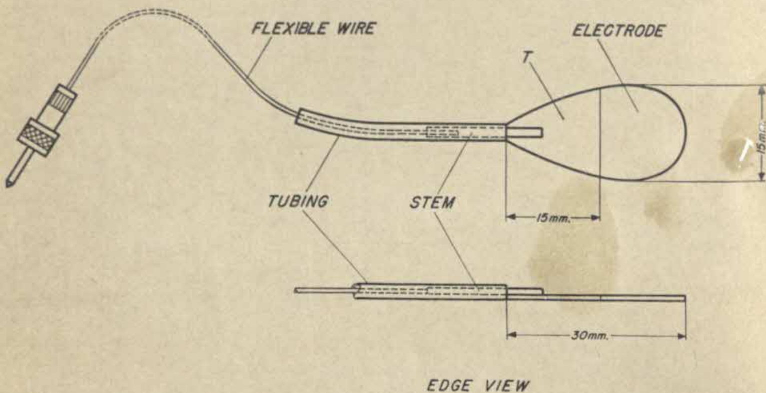


FIG. 2. DIAGRAMMATIC SKETCH OF *PC* ELECTRODE.
T represents area that is covered by Tygon primer and cement.

divided between grounds, and it is uncertain how the current will be divided.

Some of the unfavorable consequences of grounding *S* at two points on his body are as follows. The difference in resistance obtained from reversing the leads to the bridge (forward-reverse difference) is increased over the single-ground; voltage measurements on the body are disturbed, and *EKG* artifacts may appear in *EMG* tracings that had previously been free of them. Although we employ two ground electrodes, the body may be considered (for all practical purposes) to be grounded at only one point because of the small distance between electrodes.

Construction of the electrodes. Fig. 2 illustrates the *PC*-electrode which is cut from 1/32 in., fine grade silver, and is then filed and polished with steel wool until smooth. The lead wire (10 ft. of Belden 8014) is soldered to a stem which, in turn, is soldered to the small end of this electrode. After the electrode is cleaned with carbon tetrachloride, the solder joint is covered with one coat of Tygon primer and three thin coats of Tygon

paint, leaving a length of 15 mm. of uncovered silver. A solderless plug is attached to the other end of the lead for insertion in the input box of the bridge.

The purpose of the stem shown in Fig. 2 is to facilitate cleaning of electrodes. It serves as a handle while holding them. To provide this stem, solid tinned copper wire (No. 18 AWG) is doubled into the shape of a U, is soldered to the silver at one end, and to the lead wire at the other end. Actually, the solder extends along the whole length of the stem to give

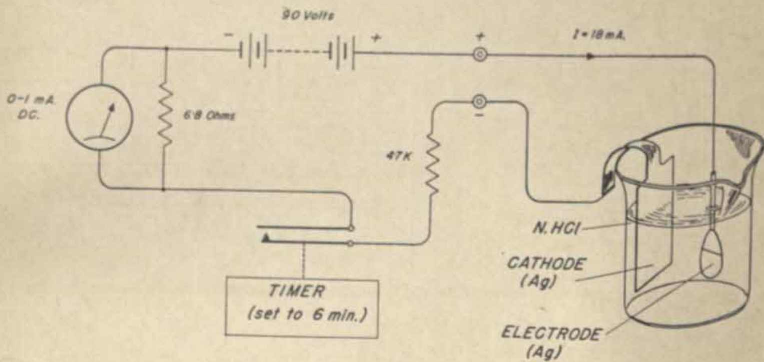


FIG. 3. CIRCUIT FOR CHLORIDING PROCESS.
See text for explanation.

stiffness to the wire. For strain relief, to prevent the wire from breaking at the point where it is soldered, a two-inch length of tubing is placed around the wire as illustrated in Fig. 2. We use Birnbach Biraco Tubing (no. 14 B & S).

Chloriding the electrodes. A diagram of the chloriding apparatus is presented in Fig. 3. The chloriding current of approximately 18 mA. is determined primarily by the value of the series resistor (4,700 Ω). Since the resistance of the chloriding bath is always a small fraction of this value, the current remains rather constant (within approximately 10%). Furthermore, the resistance of the meter-shunt is negligible as compared with the other resistors in the circuit, and hence this shunt has no effect on the operation of the circuit.

The electrode is cleaned, connected to the positive pole of the chloriding power supply, and immersed in the chloriding bath of normal hydrochloric acid. Next, the cleaned silver cathode is placed in the chloriding bath and connected to the negative pole of the chloriding power supply.

A Gralab timer is then set for a chloriding time of exactly 6 min., after which time the chlorided electrode is removed from the HCl bath, rinsed with water, and blotted dry with a soft, clean cloth. Darkening on the surface of the electrode should appear uniform. Specks on the surface of the electrode usually signify insufficient care in cleaning, and it is likely that such an electrode will not meet the voltage- and resistance-tests, and will have to be rechlorided. These tests, rather than the appearance of the electrode, determine, however, whether the electrodes may be used in

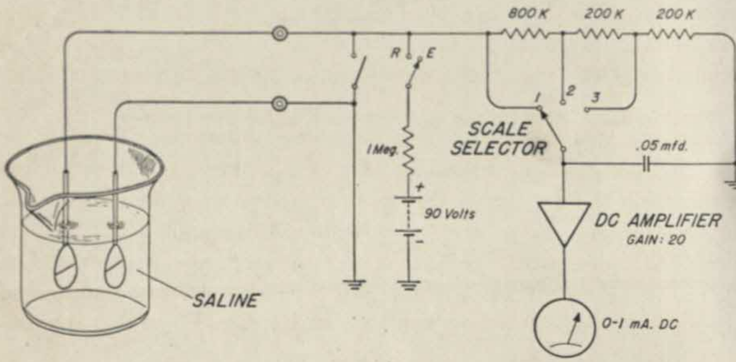


FIG. 4. CIRCUIT FOR VTVM THAT IS USED FOR CHECKING VOLTAGE AND RESISTANCE OF ELECTRODES IN SALINE.
"*DC* Amplifier" refers to direct-coupled amplifier.

Scale selection	E Millivolt range	R Range of resistance in ohms
1	0- 10	0- 100
2	0- 50	0- 500
3	0-100	0-1000

recording or must be rechlorided. If the electrode fails to meet the tests, it is best to remove the chlorided surface completely (with steel wool), and to repeat the entire chloriding process. The electrode is stored dry with protection against abrasion and mechanical damage.

Voltage measurements. Chloriding should bring the leads 'in balance,' which means that the electromotive force generated by the pair of electrodes placed in a physiological (0.9%) saline solution should be close to zero. Fig. 4 illustrates the method that we use to check electrodes for voltage and resistance. The *DC* amplifier for the vacuum tube voltmeter (VTVM) may be of low gain (*e.g.* 20) because it is used with a 1- mA. *DC* meter instead of with a recording galvanometer. The amplifier, how-

ever, should be so designed that the grid current is close to zero at the input. In practice, we have found that it is very desirable to have two *DC* amplifiers, one for measuring *PC* (illustrated in Fig. 1), and one for the purpose of taking measurements from the electrodes in saline (illustrated in Fig. 4).

One should strive to produce electrodes with no more than 1 mv. of unbalance. This requirement may seem over-strict since 1 mv. represents only 10 Ω in this circuit. Experience has demonstrated, however, that with an initial reading as low as 1 mv., the balance will usually hold during the session, whereas with higher initial readings (of the order of 5 mv.), electrode voltage may drift up to 15 mv. or more during the session, thus introducing a progressive voltage error and possible electrode-resistance increase.

Resistance measurements. After the voltage measurement, using the same electrode-checking instrument (Fig. 4), a current is passed through the electrodes and the ohm-scale is read for the two resistance values of the pair of leads, plugged in first left-to-right, and then right-to-left. A good pair of electrodes usually yields a resistance value in the neighborhood of 100-150 Ω. Reversing the electrodes will produce a different reading, but with good electrodes, the two resistance values should be within approximately 50 Ω of each other.

Electrode placement: Active lead. The electrode placed on the palm of the hand (or sole of the foot) is the active lead, and care must therefore be taken to avoid abrading the skin in the area where this lead will be placed. Gentle stroking with absorbent cotton saturated with ether cleans effectively without abrading the skin. A cellulose sponge (a block, 1 in. × ¾ in. × ⅝ in.) serves to hold the electrode (in a slot that is cut in the sponge). After saturating the sponge with physiological saline (not the more concentrated conducting jelly), the sponge is placed on the *S*'s palm, is covered with a thin piece of rubber sheeting to prevent loss of fluid, and sponge and sheeting are held securely in place by a band of elastic material (Lastonet). It is a wise practice to apply the saturated sponge to the palm as early as possible in the session that physical saturation of the palmar tissue with saline may reach its peak before the recording session commences, thus eliminating artifacts produced by the gradual soaking of the skin with saline. That is, by allowing the saline from the sponge to soak into the palm during the remainder of the preparation, any fall in resistance due merely to penetration of the salt solution into the skin should be practically complete by the time the experimental session commences.

Reference leads. In treating the skin for placement of the two reference

electrodes (see R_1 and R_2 in Fig. 1), there are two chief goals: (a) to bring resistance down as low as possible with no more than minor discomfort for *S* during the process, and (b) to bring the resistance of these leads to the same constant value in each *S*. One purpose of (a) is to prevent 60-cycle interference from various sources. In the absence of such a low-resistance ground lead, it is often impossible to secure artifact-free *EMG*s at high gain, and sometimes the 60-cycle artifact is present even in the *PC* tracing. Another purpose of bringing the reference resistance down to a low value is to minimize the contribution from reference areas to the total recorded resistance. The purpose of (b) is to minimize the undesired variation due to individual differences in resistance of the skin lying under reference electrodes. Experience in our laboratory has demonstrated that these two goals may be achieved by treating the skin of the forearm in such a way as to bring the contribution of the two reference leads, when connected in parallel, down to approximately 500 Ω. For example, with a total resistance of 10,000 Ω, approximately 9,500 Ω may be attributed to the palmar resistance and only about 500 Ω to the reference resistance.

In practice, the skin is treated in such a way as to yield a reading of 2,000 Ω (1,000 Ω from each electrode) between the two reference electrodes when they are plugged into the ohmmeter together (*i.e.* not in parallel). The final value of 500 Ω is approximated by connecting the two 1000-Ω resistances in parallel. To produce this value (approximately 500 Ω), it is essential to strive for equality in resistance from one reference lead to the other. In practice, this means that during the time that the skin is being treated, each reference electrode should be plugged in the ohmmeter with a third electrode from time to time in order to ensure that the resistance value for reference-1 to the third electrode is equivalent to the value for reference-2 to the third electrode at all stages of the resistance-reduction procedure. Unless shifts in baseline of palmar conductance are very rapid, the palmar electrode is a convenient one to employ as the third electrode.

The ventral (relatively hairless) surface of the left arm accommodates two electrodes nicely, leaving the right arm free for performance. Reference electrodes may be placed one on each arm, but this placment introduces a large *EKG*-artifact in *EMG*s from the left arm. The two reference electrodes are placed about midway between wrist and elbow on the left arm, spacing them at least 2 in. apart (and at least the same distance from other leads).

Treatment of the skin on which the reference leads are to be placed

proceeds as follows: Absorbent cotton, saturated with a grease-dissolving detergent, such as 'phisoderm,' is used to clean the skin and to reduce its resistance. Rubbing with the saturated cotton for about a minute and a half is advised. After thorough rubbing and cleaning, the electrodes are placed in the same way as that described for the palmar electrode.

At this point in the procedure, each reference electrode in turn is plugged in the ohmmeter with the third electrode in order to determine whether one reference electrode has an appreciably higher resistance than the other. In that event, further treatment is applied first to the skin under the high-resistance electrode. This treatment consists in careful application of the pointed tip of a wooden applicator (a thin wooden stick, 1-2 mm. in diameter) to the skin immediately under each electrode. The end of the applicator should be sharpened to a point, and this point dipped in electrode jelly before being applied to the skin. Although it is sometimes difficult to bring the resistance between reference electrodes down to 2,000 Ω, experience has shown that with skill and persistence, this may be achieved without vigorous scraping or injury of the skin, and with relatively little discomfort to *S*.[2]

In a recent experimental series, it was found that even after intervals of 2 hr. or more between pre- and post-calibration checks, the resistances of reference leads held very close to original values. Such findings indicate that purely physical factors such as evaporation of the saline solution in the sponge (that might increase resistance) or deeper penetration of saline into the skin (that might lower resistance) were well controlled.

---

[2] Though we do not claim that the present monopolar method is necessarily superior to bipolar methods of recording *PC*, the success of the monopolar method in the studies cited encourages us to believe that it represents an advance over methods that we had tried previously. The work of one of us (J.F.D.) with measurement of DC potentials in the operating room has led to findings that have accelerated the present application of the monopolar method to *PC* recording.

A SHUTTLE BOX FOR FISH AND A CONTROL CIRCUIT OF GENERAL APPLICABILITY

By J. L. HORNER, NICHOLAS LONGO, and M. E. BITTERMAN, Bryn Mawr College

We describe here an experimental situation designed for the study of escape and avoidance in the fish. Patterned after the situation developed for the rat by Warner, it requires the animal to shuttle back and forth from one end of an enclosure to another to escape or to avoid shock.[1] We describe here also the essential features of a control circuit designed to monitor the behavior of a group of four *S*s tested simultaneously in a set of four such shuttle boxes and to program a variety of standard training schedules. This circuit, which is quite general in function, can be used to control a set of shuttle boxes built for rats, or for insects, or for any of a number of other species. It was constructed in contemplation of a program of comparative research on escape and avoidance in a wide range of organisms.

Shuttle box. The situation used for the fish is illustrated in Fig. 1.

The animal's compartment, 10 in. long, 4.5 in. wide, and 4.25 in. high, is made of ¼-in. plastic. It is divided into two parts by a hurdle which is 1.75 in. in height. (The exact dimensions of the compartment may, of course, be varied in accordance with the size of the *S*s studied.) The floor of the compartment is opaque, and so also are the sides, except for the points at which the light-beams which actuate the photocells enter and leave. Stainless steel electrodes set into the long walls of the compartment have small openings cut in them to permit passage of the light-beams from sources to photocells. The ends of the compartment are of clear plastic, permitting illumination of the compartment by the 15-w. *CS*-lamps, which are enclosed in otherwise light-proof boxes. Since the hurdle itself is opaque, there is a marked difference in the level of illumination from one side of the hurdle to the other when one of the *CS*-lamps is turned on. The cover of the compartment is made of ¼-in. ruby plastic which permits *E* to see *S* but which, when the experimental room is darkened, eliminates visual cues from the outside. The level of water is such as to permit *S* to shuttle back and forth without difficulty but to prevent loitering in the region between the light-beams.

In escape-training, shock is turned on without warning at irregular intervals, and

* This work was supported in part by Contract Nonr 2829(01) with the Office of Naval Research and in part by Grant M-2857 from The National Institute of Mental Health.

[1] L. H. Warner, The association span of the white rat, *J. genet. Psychol.,* 41, 1932, 57-89.

S may turn it off by crossing the hurdle and breaking the beam on the opposite side. Only one of the two photocell-circuits—the one on the side opposite *S*—is functional at any given time; when its beam is broken, the other circuit becomes functional and remains so until its beam is broken, whereupon the first becomes functional again, and so forth.

In conventional avoidance-training, the onset of shock is preceded by a signal, the response to which forestalls shock. Either both *CS*-lamps are normally off and the onset of one of them serves as the signal, or both lamps are normally on and the offset of one of them serves as the signal; the lamp whose onset or offset serves as the signal may be on the same side as *S* or on the opposite side. Under what we

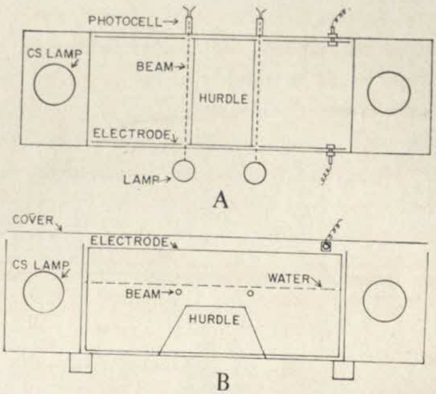


FIG. 1. THE SHUTTLE-BOX
(A = View from above; B = Side view.)

call the *Warner* condition, the shock, once it has been turned on, remains on until *S* responds.[2] Under what we call the *Hunter* condition, shock terminates after a preset interval if there has been no response during the interval; in the limiting case, the scheduled duration of shock is so brief that it terminates before the response can get under way.[3]

Our apparatus is useful also for another kind of avoidance-training, developed recently by Sidman, in which no exteroceptive warning signal is used; a recycling timer schedules shock *N*-sec. after the last response, and each response resets the timer.[4] Under what we call the *Sidman-W* condition (*W* for *Warner*), the shock remains on until it is terminated by *S*'s response. Under what we call the *Sidman-H* condition (*H* for *Hunter*), the shock terminates independently of *S*'s response. It is possible, too, to incorporate in the conventional avoidance-procedures the feature of regularly scheduled trials which are postponed by interval-response, thus con-

[2] Warner, *op. cit.*, 69.

[3] W. S. Hunter, Conditioning and extinction in the rat, *Brit. J. Psychol.*, 26, 1935, 135-148.

[4] M. Sidman, Avoidance conditioning with brief shock and no exteroceptive warning signal, *Science*, 118, 1953, 157-158.

stituting what we call the *Warner-S* and *Hunter-S* conditions (*S* for *Sidman*).

The escape-procedure and the six different avoidance-procedures already noted are generated from three procedural dichotomies. The first is *warning signal vs. no warning signal.* The second is *irregularly and independently scheduled trials vs. regularly scheduled trials postponed by interval-responses*—a *trial* being defined by the onset of the warning signal or, where no signal is used, by the onset of shock. This dichotomy really is a double one, which confounds regularity of scheduling with postponement by interval-response, but the effort required to separate the two factors in the design of a control circuit would be wasted, because the new experimental conditions which separation would permit are not very interesting. As matters now stand, at any rate, regularly scheduled independent trials are possible, but irregularly scheduled trials postponed by response are not. The third dichotomy is *shock terminated by* S's *response vs. shock terminating independently.*

In Table I, the eight experimental conditions generated by the three dichotomies are shown in their interrelations, along with the names we have given them. It may be seen from the table, for example, that the conventional *Warner* condition involves a warning signal, irregularly and independently scheduled trials, and the termination of shock by *S*'s response. *Warner-S* differs from *Warner* only in that it involves regularly scheduled trials postponed by response, and *Sidman-W* differs from *Warner-S* only in that it provides no warning signal. *Escape* differs from *Sidman-W* only in that it involves irregularly and independently scheduled trials. One

TABLE I

EIGHT TRAINING CONDITIONS GENERATED FROM THREE PROCEDURAL DICHOTOMIES

Dichotomy 1	Dichotomy 2	Dichotomy 3	Condition
(a) Warning signal	(a) Trials irregularly and independently scheduled	(a) Shock terminated by response	(1) *Warner*
(a) Warning signal	(a) Trials irregularly and independently scheduled	(b) Shock terminates independently	(2) *Hunter*
(a) Warning signal	(b) Regularly scheduled trials postponed by response	(a) Shock terminated by response	(3) *Warner-S*
(a) Warning signal	(b) Regularly scheduled trials postponed by response	(b) Shock terminates independently	(4) *Hunter-S*
(b) No signal	(a) Trials irregularly and independently scheduled	(a) Shock terminated by response	(5) *Escape*
(b) No signal	(a) Trials irregularly and independently scheduled	(b) Shock terminates independently	(6) —
(b) No signal	(b) Regularly scheduled trials postponed by response	(a) Shock terminated by response	(7) *Sidman-W*
(b) No signal	(b) Regularly scheduled trials postponed by response	(b) Shock terminates independently	(8) *Sidman-H*

of the eight conditions shown in the table—the sixth—is meaningless; permitting neither escape nor avoidance, it involves a series of unheralded shocks about which *S* can do nothing.

Control-circuits. The control-circuits for the four shuttle boxes are constructed on four identical boards which, with their associated timing clocks, are mounted on a standard, 6-ft., steel relay rack. The boards are connected to the shuttle boxes by octal plugs and sockets, which makes it easy to substitute shuttleboxes suitable for other animals. A master control-panel, which holds a latency-measuring clock, a shock-interrupting circuit, and a set of selector switches, also is mounted on the rack. One of the individual circuits is shown in Fig. 2 with the functions of the master indicated. Switches on the master-panel permit *E* to select one of three basic modes of operation: the *manual Warner-Hunter,* the *automatic Warner-Hunter,* and the *Sidman.*

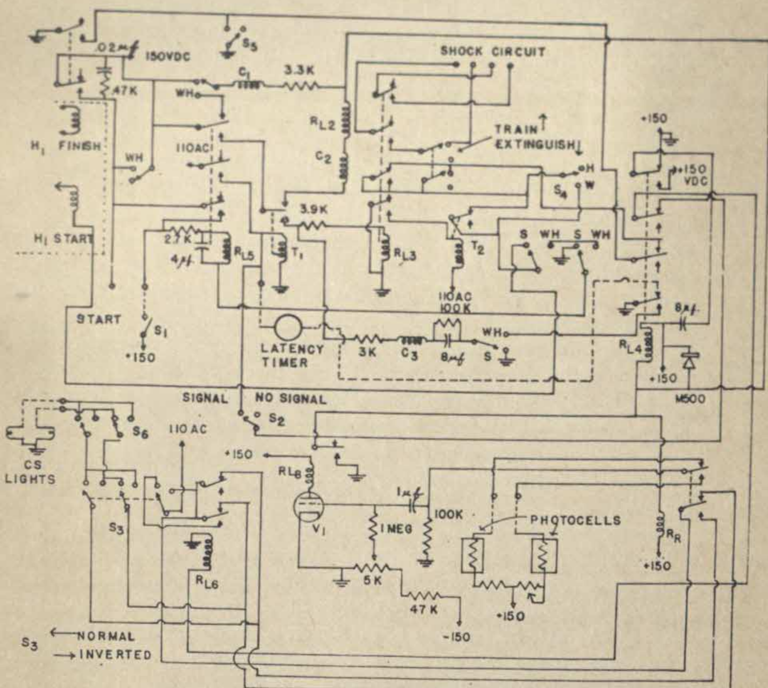


FIG. 2. CONTROL CIRCUIT FOR ONE OF THE SHUTTLE-BOXES
(The functions of the master are indicated by broken lines. V_1 is a 12AT7.)

In the manual Warner-Hunter mode, conventional avoidance training (Conditions 1 and 2) and escape-training (Condition 5) are studied. In these conditions, trials are irregularly and independently scheduled. The separate control-circuits for the four shuttle boxes simultaneously keep track of *S*'s position and count interval-responses (C_1), but *E* manually programs the trials to one box at a time and reads the latency of response from a clock. A selector-switch determines which box is to have a trial, and the trial is initiated by the depression of another switch, S_1, which energizes RL_5. This relay starts the *CS*-timer, T_1, and the latency-clock, and it activates RL_6, which initiates the *CS*. The purpose of RL_6 is to permit choice between *warning signal* and *no warning signal* (S_2), and between *normal* and *inverted* operation (S_3). In normal operation, both *CS*-lamps are off, and the signal, if any, is the onset of one of them; in inverted operation, both lamps are on, and the signal is the offset of one of them. The position of S_6 determines whether the signal will occur on the side of *S* or on the opposite side. In escape-training, the *CS*-timer is set at zero, and there is, of course, no signal.

When the *CS*-timer has run a preset interval, it energizes RL_3, which connects the shocking electrodes to the shock-supply and turns on the shock-interrupter if it is desired (otherwise the shock is continuous). RL_3 also energizes the shock-timer, T_2. In Condition 2 (selected by Switch S_4 in Position *H*), this timer terminates the trial by turning off RL_5 after it has run a preset interval. In Conditions 1 and 5 (Position *W*), the trial is terminated only by response, which fires the response-relay, RL_8, in the plate-circuit of the photocell amplifier, V_1. RL_8 fires RL_4 which turns off RL_5 and steps the ratchet-relay which shifts photocells and *CS*-lamps. Since the operation of RL_8 is very rapid, RL_4 switches an 8-μf. capacitor across its coil to give a holding time of about 1/4 sec. In extinction, RL_3 turns off RL_6. The activation of RL_4 stops the latency-clock, which *E* reads and resets.

In the automatic Warner-Hunter mode, trials are initiated in the four shuttle boxes simultaneously by a relay in the master which is actuated by a film-timer. The latency-clock is disconnected, and data are taken by counters C_2 and C_3, which record avoidances and escapes, respectively. As in the manual mode, C_1 records interval-responses. The counters, of course, give only the number of responses of each kind occurring in a given time, which is all that may be necessary. When the precise distribution of avoidances over a series of trials must be known, a continuous-feed kymograph is used, with one pen to show the avoidances in each box and a fifth pen to mark off the trials. The apparatus functions unattended in this mode, which is more efficient than the manual mode for conventional avoidance-training (Conditions 1 and 2) and for escape-training (Condition 5) when response-latencies are not required.

In the Sidman mode, also automatic, Conditions 3, 4, 7, and 8 are studied. The apparatus functions as it does in the automatic Warner-Hunter mode, except that trials are programmed separately for each box by a Hunter timer, H_1, which is reset by response when the multigang switch, *WH-S* in Fig. 2, is in the Sidman position. The same switch changes the functions of the counters. In this mode, C_1 counts all responses; C_2 counts responses in the last *x* seconds before shock is due (an interval during which, in Conditions 3 and 4, a signal is presented); and C_3 counts the total number of shocks. Where information as to the temporal course of responding is required, a cumulative response-recorder is so connected to the control-circuit that

the stepping mechanism is actuated by every response, one event marker by response in the last x seconds, and another marker by shock-onset.

There are several extinction procedures of interest in relation to Sidman training, two of them particularly appropriate to Conditions 7 and 8. One is simply to eliminate shock, which is accomplished by turning off H_1. A second is to schedule regularly recurring shocks independently of the animal's behavior; when S_5 is open, H_1 is not reset by response (that is, shock is not postponed), although, when S_4 is the position W, shock remains on (once it has come on) until terminated by response. Two further extinction procedures are appropriate to Conditions 3 and 4. In the first, interval-response resets H_1, thus postponing the signal or terminating it if it has already begun, but there is no shock for failure of response to the signal; with the shocking voltage reduced to zero, S_4 in position H, and the shock-timer set for a fraction of a second, H_1 is reset without shock at the end of the *CS*-interval. In the second, accomplished by opening S_5, response to the signal resets H_1, but

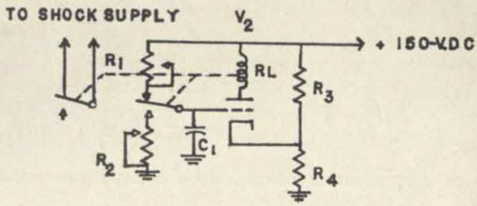


FIG. 3. SHOCK-INTERRUPTING CIRCUIT
(R_1, R_2, 250K; R_3, 10K, 5 w., wire-wound; R_4, 1K, ½ w., carbon; V_1, 6C4; C_1, 1.0 μf., 400 v.; *RL*, Potter and Brumfield No. MH11LM.)

interval-response does not; whether the animal is shocked for failure of response to the signal depends on whether the shocking voltage is reduced to zero.

There is liberal use of arc-suppression in this apparatus to protect relay-contacts and thus to insure reliability of operation. The diode which protects the contacts of the response-relay, R_8, is an M500. A ratchet-relay of good quality (*e.g.* Potter and Brumfield No. PC11A) is required, with gold-flashed contacts and a separate dust-cover, because the photocell-circuits are very 'dry' and may be prevented from operating by a small bit of dust or tarnish on the contacts. The *CS*-timer and the shock-timer are of the stalled motor-clutch variety and reset with the removal of voltage.

Shock-interrupter. A switch in the master control-panel enables E to choose one of two shocking circuits—a constant-voltage supply for the fish and other aquatic forms, and a constant-current supply for non-aquatic forms. Since continuous shock at a level high enough to be effective tends to tetanize the fish and certain other animals, we use the shock-interrupter shown in Fig. 3.

The circuit is fairly simple, but it was hit upon only after some false starts with thyratrons. R_3, R_4 is a voltage-divider which keeps the relay, *RL*, in the plate-circuit from being energized. (In the interests of reliability, *RL* is a 10,000 Ω,

sealed relay.) The novel feature of this circuit is that the on-time and the off-time can be set independently. For the present purpose, R_1 is a fixed resistance which gives a 0.25-sec. on-time; the off-time is varied by adjusting R_2. The x in the B+ line shows the place at which the circuit is energized by RL_4 of Fig. 2; since the

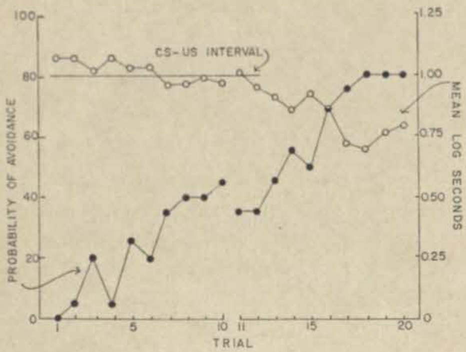


FIG. 4. THE PERFORMANCE OF TWENTY GOLDFISH IN A CONVENTIONAL AVOIDANCE PROBLEM
(Ten trials per day were given with a *CS–US* interval of 10 sec. and an average intertrial interval of 3 min.)

contacts of *RL* in series with the electrodes are normally closed, the animal is shocked as soon as RL_4 is energized.

Sample results obtained with this apparatus are presented in Fig. 4. The curves show mean log latency of response and probability of avoidance for 20 goldfish given 20 trials of training under Condition 1 (Warner) with a 10-sec. *CS–US* interval. The apparatus functioned in the manual Warner-Hunter mode.

THE FORCE-PLATFORM: INSTRUMENT FOR SELECTING AND TRAINING EMPLOYEES

By JAMES W. BARANY, and JAMES H. GREENE, Purdue University

In 1949, a French engineer, Lucien Lauru, developed an instrument which he called the "Effort-detector."[1] This device utilized the unique property of pizo-electric crystals which enable them to emit a small electric impulse that is proportionate to the amount of externally applied pressure. Five such crystals were incorporated into a triangular platform in such a manner that they would sense the orthogonal components of the forces exerted by an operator whenever a body-movement is made while performing on the apparatus. After proper amplifications, the emitted electric impulses are recorded to produce a graphic picture of the forces in the vertical, horizontal, and transverse planes required to execute a specific movement.

A force-platform designed and constructed recently under the auspices of the Purdue Farm Cardiac Project (Fig. 1) has revealed some interesting characteristics.[2] The device is a highly sensitive instrument for the measurement of forces exerted by an operator over a period of time. In fact, the apparatus is so sensitive that it can detect the heart beat of *S* standing on the platform. The purpose of this paper is to discuss some of the practical problems in applied psychology that might be solved by this device.

Muscular skill can be defined as the ability to regulate the degree of force and the direction of force exerted during an operation.[3] A highly skilled individual makes different movements, or the same movements, more quickly and exhibits fewer stops or hesitations than a less skilled operator. The reason some operators perform better than others is their ability to establish a rhythmic pattern with few extra or false motions. Production personnel for many jobs should be selected on the basis of coordination of gross body-movements and their sense of timing or rhythm. The force-platform would be used as a personnel-selection tool for predicting future success on jobs requiring rhythmic patterns of motion.

Another area of application would be in the field of personnel training based upon

[1] Lucien Lauru, Physiological study of motions, *Advanced Management,* 22, 1957, 17-24.

[2] J. H. Greene and W. H. M. Morris, The force platform: An industrial engineering tool, *J. Ind. Engr.,* 9, 1958; 128-132. The design of a force platform for work measurement, *J. Ind. Engr.,* 10, 1959, 312-317.

[3] Abraham Cohen, *Time-Study and Common Sense,* 1947, 26.

the principles of movement-analysis as defined by Lindahl.[4] This method has been successfully employed in the training of operators for cutting small tungsten disks on a machine using a rotary, abrasive, cutting wheel. Lindahl constructed a simple graphic recording device to establish the stroke pattern of the abrasive wheel. These graphic representations of the cutting cycles were shown to the trainees to point out specific errors in their performance. Unfortunately, the physical design of the machinery involved in many industrial situations makes it impossible to adopt this procedure.

Fig. 2 gives some indication of how the force-platform could be used for selection and training employees. The graphs show the frontal forces exerted by 3 *S*s (men)

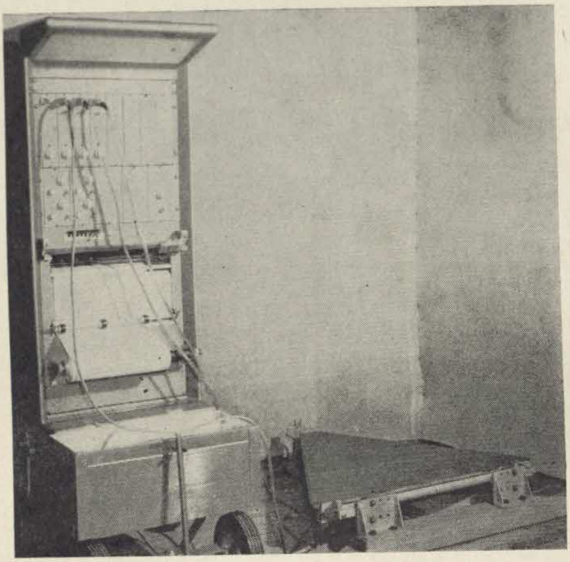


FIG. 1. PURDUE FORCE-PLATFORM

performing a simple experiment with a foot-lever. The operator had to push with a 35-lb. force at a rate of 44 strokes per min. S_1 was in excellent physical condition and had spent most of his life on a farm doing physical work including the operation of tractors equipped with foot-pedals similar to those used in the experiment; S_2 was an athlete but had never worked extensively on a job requiring muscular exertion; and S_3 had little contact with physical activities either in his vocation or avocation.

It is quite apparent that S_1 pushed the foot-pedal in one quick surge of power, while S_2 took a little longer, and S_3 took still longer. The results were that S_1 did his task in one movement and rested between strokes while, at the other extreme, S_3 spent most of the cycle time with his muscle contracted against the foot-lever. This

[4] L. G. Lindahl, Movement analysis as an industrial training method, *J. appl. Psychol.*, 1945, 29, 420-436.

gross example does indicate how the force-platform could be used for personnel selection. A person familiar with reading these diagrams can detect even minor variations in the tracings.

The platform could be installed next to a machine to record the movements of the operator while performing the operation. Its greatest value, however, lies in formulating programs in vestibular training. Here a simplified mock-up of the operation could be constructed around the force-platform. New trainees would be able to have a continuous record of their progress and could see graphically the target, the pattern

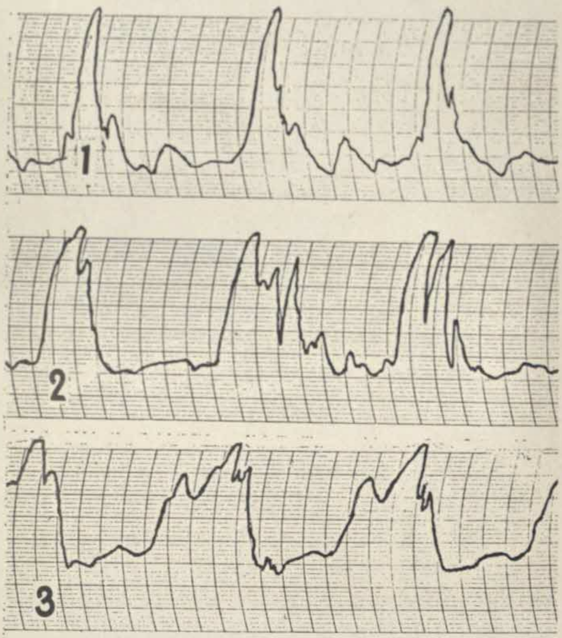


FIG. 2. GRAPHIC RECORDS FROM THE FORCE-PLATFORM

(1) Record of S_1, age 22 yr., height 6 ft., weight 175 lb., medium build, athlete and farm hand; (2) Record of S_2, age 22 yr., height 5 ft. 11 in., weight 170 lb., medium build, athlete; (3) Record of S_3, age 22 yr., height 5 ft. 9 in., weight 155 lb., medium build, nonathletic.

of motion, they are trying to achieve. Progress could be easily measured by placing the trainee's force-pattern next to the pattern established by the skilled operator.

The future developments of the force-platform could be in the area of applied psychology for the purpose of selecting and training employees. This device, which detects and transcribes the movements of an operator, could be used to determine the ideal work-patterns and then to show the worker how to do the task in the best possible way.

The apparatus may also be employed in pure research. For example: the force-platform could be used to study the onset of physiological fatigue with its accompanying loss of coördination; for experiments concerned with the acquisition of motor skills; as a technique for checking the various hypotheses resulting from different theories on learning; for analysis of movements in reaction-time studies; and perhaps it could even be used as a means of qualtifying the activity-level of a client in an interview-situation. In fact, a miniature model of the platform could be used in comparative psychology in order to measure the reaction of a rodent to electrical shock or some other noxious stimulus. Undoubtedly, the research oriented psychologist can think of many other possible applications of this instrument.[5]

---

[5] The authors acknowledge the assistance of Dr. W. H. M. Morris, Associate Professor of Agricultural Economics, in designing the platform.

A SIMPLIFIED ELECTRODE-ASSEMBLY FOR IMPLANTING CHRONIC ELECTRODES IN THE BRAINS OF SMALL ANIMALS

By ELLIOT S. VALENSTEIN and WILLIAM HODOS, Walter Reed Institute of Research, and LARRY STEIN, Wyeth Institute for Medical Research

Several techniques have been described for implanting multiple electrodes in the brains of large laboratory animals.[1] With smaller animals, such as the rat and guinea pig, it is more difficult due to the limitations of space and to the fragility of the skull. For the past two years, we have used a method which enables us to place at least four electrodes in the rat's brain. Many of our *S*s have been tested daily for periods in excess of one year, and the method has proven to be completely reliable. All components are easily accessible, and the assembly has the advantage of simplicity of construction and economy. The technique has proven to be equally useful for stimulation and for recording, and it is applicable to a wide variety of laboratory animals.

Construction of electrodes. The electrode-pedestal and the male-connector both are constructed from size 6-32 threaded nylon rods which can be purchased in convenient lengths.[2] The nylon may be cut with a scalpel blade into approximately 5/16-in. segments. With larger animals, it is desirable to cut the nylon in ½-in. lengths. Two longitudinal holes, 0.032-in. (drill bit #67) in diameter, with centers 0.045 in. apart, are drilled through the nylon rods. It is desirable to place these holes at equal distances from the center, which can best be accomplished with the aid of a simple jig that holds the nylon rod and guides the drill bits.

The point of a 21-gauge hypodermic needle is cut off with a fine emery wheel and then forced into one of the holes until approximately ¼ in. extends out the opposite side. We have found 0.010-in. stainless steel wire,

---

[1] J. M. Burlo, A multiple electrode for depth recording, *EEG Clin. Neurophysiol.*, 7, 1955, 655; J. M. R. Delgado, Evaluation of permanent implantation of electrodes within the brain, *EGG Clin. Neurophysiol.*, 7, 1955, 637-644; G. C. Sheatz, Multilead electrode holders in chronic preparations, In D. E. Sheer (ed.), *Electrical Stimulation of the Brain,* in press.

[2] The threaded nylon rods and nuts are manufactured by the Weckesser Company, 5701 Northwest Highway, Chicago 30, Illinois.

which can be obtained quadruple insulated from the factory,[3] very convenient for macro-stimulation and recording. Approximately ⅛ in. of insulation is scraped from a section of wire 5-6 in. long; the bared wire is inserted into the hypodermic needle and crimped with a pair of blunted diagnoal wire cutters (Fig. 1 A). The hypodermic needle now is pulled back until the cut end is flush with the bottom end of the nylon rod. The emery wheel is used for cutting the needle flush with the top end of the rod (Fig. 1 B). This procedure is repeated on the other side. The ends of the electrode-wire are inserted in a drill press and, while the nylon pedestal

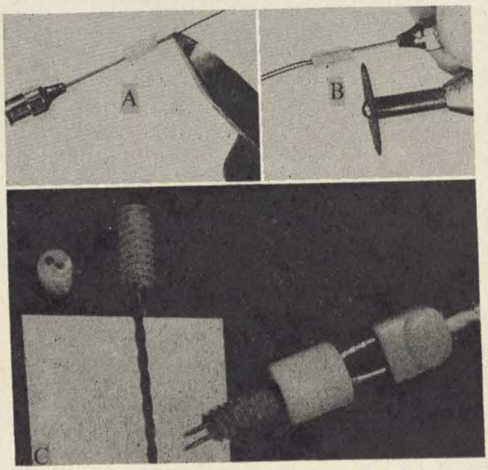


FIG. 1. VARIOUS STAGES IN THE CONSTRUCTION OF THE ELECTRODE-ASSEMBLY AND MALE-CONNECTOR

(A) Hypodermic needle has been forced through hole in pedestal and is shown being crimped to electrode wire. (B) Two electrodes are shown emerging from bottom end of pedestal. Emery wheel is used to cut hypodermic needle flush with top of pedestal. (C) Side and top view of finished electrode, and side view of male-connector and threaded, locking collar.

is held, the wire is braided at a slow speed. Fig. 1 C illustrates the braided electrode wire. The electrode is cut to a convenient length for storing, then dipped in a vinyl lacquer up to the base of the pedestal, and left to dry for 24 hr.

The male-connectors are made essentially the same way, except that 0.018-in. steel, spring wire (piano-wire) is used instead of the electrode-wire. The piano-wire should extend approximately 3/16 in. beyond the

[3] Type 316 stainless steel wire (quadruple insulated) may be obtained in sizes 0.003, 0.005, 0.007 and 0.010 in. from the Driver-Harris Company, Harrison, New Jersey.

nylon rod. A threaded collar, which is used to tie the connector to the electrode, can be made easily from a nylon rod, or nylon nuts can be purchased with the threaded nylon rods. Fig. 1 C illustrates a top and side view of the electrode-assembly and a side view of the male-connector. Hearing-aid cords, which can withstand considerable twisting without breaking, have been found useful for work with unrestrained animals.

Implantation. Prior to implantation, the electrode must be cut to proper length. It is desirable to have the base of the pedestal resting either directly on the skull or so close to it that it does not extend too high. When bony landmarks are used for stereotaxic implantation, it is easy to esti-

FIG. 2. A RAT'S SKULL PREPARED FOR IMPLANTATION OF ELECTRODES.

mate the length of the electrode. With monkeys and cats, it is useful to have several skulls available. The electrode can be cut with sharp wire-cutters, and, if desired, the enamel may be scraped back farther with a scalpel blade; with a little care it is not difficult to cut each member of the pair to a different length for recording or stimulating between points some distance apart.

Fig. 2 illustrates the skull of a rat which has been prepared for implantation. The electrode can be held conveniently in a rod that has a 6-32 internal thread. Three stainless steel screws, size 0-80, with round heads (1/8 in. long for rodents; 3/16 in. for larger animals), have been screwed into the skull. When the skull has previously been drilled with a #56 bit, these screws will tap themselves into the skull and anchor securely with a few turns. The screws help to bind the cement to the skull. Shallow holes also are drilled into the skull to serve as additional gripping points for cement. The electrode then is inserted, and an acrylic cement is placed over the skull, the screws, and half way up the pedestal. We have found cranioplastic cement to be very desirable for this purpose, since it

generates relatively little heat and makes an excellent bond with skull, screws, and nylon pedestal.[4] As this cement requires 15-20 min. to set completely, there is sufficient time to mold the cement smoothly over the skull. The threads on the nylon pedestal serve as a means of locking the connector and electrode-assembly together (Fig. 3), not for threading into the skull during implantation. If more than one electrode is to be inserted, a quick-drying cement can be used to anchor all but the last electrode, and the cranioplastic material used for the final molding. If the skin-incision is made to one side, the pedestals may be inserted through stab-wounds in the skin, and suturing can be accomplished at a distance from the pedestals.

FIG. 3. VIEW OF A RAT WITH THREE ELECTRODES ONE MONTH FOLLOWING IMPLANTATION
(The male-connector has been plugged into the posterior electrode.)

Fig. 3 illustrates a rat with three pairs of electrodes, one of which has the male-connector in place. With this technique, it is possible to place four or five chronic electrodes in a rat's brain, provided that no pedestal be closer than 4.5 mm. to its neighbor.

Indifferent electrodes for unipolar stimulation and recording also have been used with this technique by simply substituting a bared wire for one of the electrode-wires and placing the former in a tight coil on top of the skull under a convenient muscle. With little modification, the technique may be used with any size or type of wire preferred. The only change necessary is to substitute the appropriate size hypodermic needle, spring wire, and hole through the nylon pedestal.

[4] The cranioplastic cement is manufactured by the L. D. Caulk Company, Milford, Delaware and is distributed as a Cranioplastic Kit (N-196) by Codman and Shurtleff, Inc., 104 Brookline Avenue, Boston 15, Massachusetts.

APPARATUS NOTES

A VIDEO-MOTOR BOARD

Many pegboards currently in use give adequate measurement of finger dexterity or of eye-hand coördination on a small work surface. In many cases, however, gross arm movements are more predictive of on-the-job performance than are fine finger movements. Size of a motor skills pegboard, therefore, is a very important aspect of its construction. The Shurrager Video-Motor Board was designed to allow for a large test-area in which measurements of 'pick-up,' 'transport,' and 'release' could be obtained. Several variations of this board have been used at the psychological laboratory of this Institution and tested in industry. The latest variation is described here.

The total video-motor board is assembled from six separate components: Four die-molded, laminated, plastic boards $12\frac{3}{8} \times 8\frac{3}{8}$ in. each containing 25 holes $\frac{5}{8}$-in. square and $\frac{5}{8}$-in. deep arranged in five rows. Horizontal distance between

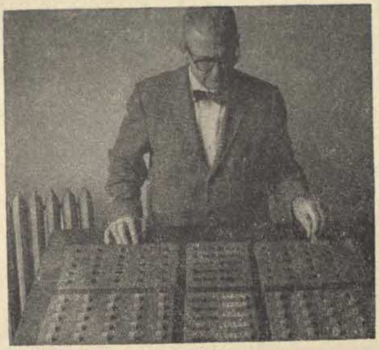


FIG. 1. THE SHURRAGER VIDEO-MOTOR BOARD

holes is $1\frac{3}{4}$ in.; vertical distance between holes is $\frac{3}{4}$ in. Two die-molded, laminated, plastic boards $8\frac{1}{2} \times 7\frac{1}{2}$ in., each containing five pairs of grooves $2\frac{5}{8}$ in. long, $\frac{5}{8}$ in. wide, and $\frac{5}{8}$ in. deep. Pegs, $\frac{1}{2}$ in. square, have been cut in lengths of $1\frac{1}{8}$-in. from $\frac{1}{2}$-in. square aluminum rods. The pegs fit loosely into the holes and grooves.

The component boards are backed with steel which makes possible the use of magnetic metal pegs or of magnetized rubber pegs where such are desired to simulate actual assembly operations or when work-surface angle is a factor. The steel back is covered with a soft rubber backing to prevent slipping. When all six boards are used, the minimal size of the apparatus is $32\frac{1}{4} \times 16\frac{3}{4}$ in. Obviously, there is no maximal size, since the parts may be separated as widely as is desired (see Fig. 1).

Time- and error-scores have been derived from the board for six elements: (1)

'Out-in'; a bimanual operation in which the pegs are moved from the sides into the middle grooves. (2) 'In-out'; a bimanual operation in which the pegs are moved from the middle grooves to the original position of Element 1. (3) 'Right-cross'; a mono-manual operation with the right hand in which pegs are moved from the lower right component to the upper left. (4) 'Left-cross'; a mono-manual operation with the left hand in which the pegs are moved from the lower left component to the upper right. (5) 'Right-back'; a mono-manual operation with the right hand in which the pegs are moved from the upper right to the lower right component. (6) 'Left-back'; a mono-manual operation with the left hand in which the pegs are moved from the upper left to the lower left component (shown in Fig. 1).

It is also possible to obtain scores from other elements such as 'Left-up,' 'Right-up,' etc.

This apparatus has been used in several studies which range from purely applied research, such as testing work-surface angles, to more theoretical studies designed to show the contribution of the kinesthesis to intelligence test-scores of human *S*s. This board has an important and often overlooked feature; namely, its adjustable size. Studies of motor learning may often be unproductive, so far as their theoretical contributions are concerned, because the movements measured are too small to allow for kinesthetic feedback.

Our original boards were of one-piece construction. The new board, with its variable size, has proved much more satisfactory.

Illinois Institute of Technology P. S. SHURRAGER

NOTES AND DISCUSSIONS

REVERSIBLE PERSPECTIVE AS A FUNCTION OF STIMULUS-INTENSITY

Some years ago, Fisichelli reported the results of a number of experiments on reversible perspective using the Lissajous figures, and advanced the theory that the rate of reversal is a function of the amount of stimulation of a given set of nerve fibers.[1] This conclusion seemed well supported by the evidence. Since this work appeared, little advance in the understanding of the factors responsible for reversible perspective has been made, but the question has developed additional interest as a result of Eysenck's finding that the reversal rate is reduced in psychotics,[2] and Speakman's that it is reduced in old people.[3] It is possible that further understanding of the processes taking place in psychosis and aging could be gained from a renewed consideration of the factors affecting rate of reversal. The implication of Fisichelli's theory, taken in conjunction with the findings on psychosis and aging, appears to be that among psychotics and old people there is reduced sensitivity to stimulation.

This argument needs strengthening at two points. In the first place, Fisichelli worked with the Lissajous figures while Eysenck used the Necker Cube and Speakman the Schroder staircase. It has been shown by Thurstone that a number of tests of the rate of reversible perspective are positively intercorrelated.[4] It seems likely, therefore, that the Lissajous figures and the two dimensional figures are comparable tests and that it is legitimate to argue from one to the other. Nevertheless, a demonstration that reversal rate on the two dimensional figures is a function of the amount of stimulation is clearly desirable. Secondly, Fisichelli did not experiment with the variable of stimulus-intensity, which might be regarded as one of the simplest and most direct estimates of the amount of stimulation of the nerve fibers. Accordingly, on the basis of Fisichelli's work two hypotheses have been set up. First, rate of reversal on the Necker Cube should be a function of amount of stimulation; and secondly, a further measure of the amount of stimulation can be derived from the stimulus-intensity,

[1] V. R. Fisichelli, Reversible perspective in Lassajous figures: Some theoretical considerations, this JOURNAL, 60, 1947, 240-249.

[2] H. J. Eysenck, *The Scientific Study of Personality,* 1952, 220.

[3] Quoted by A. T. Welford in *Ageing and Human Skill,* 1958, 170.

[4] L. L. Thurstone, *The Factorial Study of Perception,* 1944, 109.

and therefore stimulus-intensity should affect reversal rate on the Necker Cube.

The apparatus consisted of a wooden box about 30 cm. square, in one side of which was a glass window painted black. A Necker Cube measuring 4 cm. square was scratched in the paint, so that the cube, when a lighted bulb was placed in the box, was clearly observable. Two degrees of illumination were used: a 25-w. and a 60-w. bulb. Twelve men, university students, were used as subjects (*S*s). They sat about 2 m. from the box. After the nature of the reversal was explained to them, they were shown the cube until they reported reversals. When they had become familiar with their task, they were then asked to report, during a 30-sec. interval, the number of (1) spontaneous reversals and (2) of reversals that occurred when they were trying to prevent them. Half the *S*s looked at the brighter cube first; half the dimmer.

The mean reversal rates reported for the two conditions under the different stimulus-intensities were as follows: Condition 1, 25-w. light, 7.51; 60-w., 12.66 ($t = 2.64$, $p > 0.05$); Condition 2, 25-w., 3.75; 60-w., 7.25 ($t = 5.00$, $p > 0.001$). It will be observed there is a significant tendency for the *S*s to report more reversals at the higher stimulus-intensity.

The results confirm and extend Fisichelli's theory that rate of reversal is a function of the amount of stimulation by showing that reversal rate is a function of stimulus-intensity and that the relationship holds for the Necker Cube as well as for the Lissajous figures. The wider implications of this finding seem to be twofold. In the first place, the demonstration that reversal rate depends on the intensity of the stimulus relates reversible perspective to Hull's construct of stimulus-intensity dynamism (V) and thereby to a number of other phenomena. Secondly, the findings that psychotics and old people have a low reversal rate implies that psychotics and old people perceive stimuli less intensely or, in Hullian terms, have low values of V.

An alternative explanation could be found in the concept of attention. It is possible that reversals are a function of attention, and that the present findings were obtained because one determinant of attention is stimulus-intensity.[5] The implication of this explanation would be that there is a reduction in the capacity for attention among psychotics and old people. Such an explanation accords well with recent work suggesting that psychotics

---

[5] D. E. Broadbent, *Perception and Communication*, 1958, 298.

and old people have high inhibitory potentials and are impaired on tasks where sustained attention is required.[6]

University of Exeter, England RICHARD LYNN

DEPTH-PERCEPTION AND ASTIGMATISM

Depth-phenomena described here were observed and investigated by the author whose eyes are astigmatic. These phenomena result from the fact that the astigmatism differs in kind and degree for each eye. When I look at a dot from a distance of 30 cm. with my right eye, I see it as two dots—one slightly above and to the left of the other. The lower dot is 0.58 mm. (6′43″ of angular value) to the left and about 1 mm. (12′10″) below the upper one. With my left eye, I see it similarly doubled; the lower dot is seen almost vertically beneath the upper one, 0.29 mm. (3′21″) to the left and 0.89 (10′11″) below the upper one. With the right eye, the lower dot looks larger and sharper than the upper; with the left eye, the upper looks larger and sharper than the lower. When I look at the dot binocularly from the same distance (30 cm.) I see again two dots; both approximately of the same sharpness and about the same size (effect of the summation of both images), but the upper one is farther away from me and the lower, which is nearer, is a bit to the left.

When the observation is repeated with a vertical hair, or thin line drawn on a sheet of paper, similar results are obtained. With the right eye, I see two lines slightly displaced vertically and horizontally, the right line being sharp, the left thinner and blurred. With the left eye, I also see two lines, the right thin and blurred and the left sharp and clear, the lateral distance between the two lines being larger for the right eye than for the left; for both eyes the interspace between the two lines is filled up with a brownish blur. When viewing with both eyes, I see two lines displaced in space; the left one in front of and slightly lower than the right one and between them a brownish blur. Besides, I see the portion of the sheet of paper in the neighborhood of the left image nearer than that adjoining the right image, as if the sheet were broken by the two images of that line.

These results are the same as those obtained from two dots or two lines similarly arranged on the two halves of a stereogram and combined in free stereoscopy or by the use of a stereoscope.

[6] P. H. Venables and J. Tizard, Paradoxical effects in the reaction-times of schizophrenics, *J. abnorm. soc. Psychol.*, 53, 1956, 220-224; Stephen Griew and Richard Lynn, The construct "reactive inhibition" in the interpretation of age changes in performance, *Nature*, 1960, 182.

If astigmatism is corrected in one eye by use of an appropriate lens that only one line is seen by that eye, then in binocular vision the so-called Panum Effect appears, *i.e.* two lines at different distances (right line nearer if the astigmatism is compensated by glass in the left eye and vice versa); this effect, however, is not so steady as in normal arrangement for stereoscopic fusion.

My eyes show varying differences in the kind and degree of astigmatism of the two eyes at different distances and consequently depth-differences between the two lines by binocular vision also change.

The perception of depth obtained from astigmatic images confirms the laws of stereoscoping vision, especially the law of intersection of visual lines, and supports the theories of correspondence and disparity.[1]

Ottawa, Canada J. L. ZAJĄC

ERRATUM

Line 9, p. 523, in the article on "The effect of muscular involvement on sensitivity," by J. H. McFarland, Heinz Werner, and Seymour Wapner (December, 1960, Vol. 73), should read:

"visual field has been in terms of a corresponding distribution of rods and"

The line of type that appears there is Line 13 that was marked for correction in the page proof. The error in Line 13 was corrected but due to an accident occurring during this process, Line 9 fell out of place and the line of type removed from Line 13 was substituted for it—the article then being forthwith released for the presses. A curious error but one that no proofreader had an opportunity to catch.

K. M. D.

AN ACKNOWLEDGMENT

The JOURNAL is indebted to Professor Leo Postman for the portrait and signature of Edward C. Tolman reproduced in the frontispiece of this number. The portrait is a reproduction of a photograph, taken while he was sitting for the painting that was presented to the University he so long served, which was completed shortly before his death. A full signature could not be obtained as he normally abbreviated his second name.

K. M. D.

---

[1] J. L. Zając, Depth perception of stereoscopic images resulting from fusion of crossed and uncrossed double images, this JOURNAL, 72, 1959, 163-183.

Edward Chace Tolman: 1886–1959

Edward Chace Tolman was born in Newton, Massachusetts, on April 14, 1886, of a well-to-do family. As he described it, the family had a Puritan tradition of hard work and a Quaker tradition, on his mother's side, of plain living and high thinking. These values stuck with Tolman throughout his life.

His older brother, Richard, who later became a distinguished theoretical chemist and physicist at the California Institute of Technology, came to be for him an admired intellectual and scientific model and also a somewhat frightening rival. Edward Tolman followed his brother to the Massachusetts Institute of Technology and in 1911 received his B.S. degree in electrochemistry. He then entered Harvard as a graduate student in the joint Department of Philosophy and Psychology. As he later explained it, this choice of psychology over physical science as a career reflected a combination of feelings of inadequacy in mathematics and manipulative skill, a desire to study human conduct and to aid in the betterment of mankind, and a wish to avoid a field of science too closely competitive with his brother's.

After receiving his doctor's degree from Harvard in 1915, Tolman became an instructor at Northwestern University for three years. In the meantime he made a trip to Germany where he had his first exposure to Gestalt psychology which later was strongly to color his thinking.

Tolman's distinguished career as a creative scientist and as a leader in the academic and social community spanned the four decades from 1918, when he came to the University of California in Berkeley as an instructor in psychology, until his death in Berkeley on November 19, 1959. His scientific eminence was recognized in the many honors bestowed upon him. He was elected a member of the Society of Experimental Psychologists in 1930, of the National Academy of Sciences in 1937, of the American Philosophical Society in 1947, of the American Academy of Arts and Sciences in 1949; and Honorary Fellow of the British Psychological Society in 1954. He was awarded honorary degrees by Yale, McGill and the University of California. Posthumously he was elected Associate Member of the Société Française de Psychologie, and a member of the Sociedad Española de Psicologia.

He was president of the American Psychological Association in 1937, president of the Society for the Psychological Study of Social Issues in 1940, vice-president of the American Association for the Advancement of Science in 1944, president of the APA's Division of Theoretical-Experimental Psychology in 1946 and of its Division of General Psychology in 1953, and co-president of the XIVth International Congress of Psychol-

ogy in 1954. In 1957, he received an award for Distinguished Scientific Contribution from the APA, with the following citation:

> For the creative and sustained pursuit of a theoretical integration of the multifaceted data of psychology, not just its more circumscribed and amenable aspects; for forcing theorizing out of the mechanical and peripheral into the center of psychology without the loss of objectivity and discipline; for returning man to psychology by insisting upon molar behaviorism purposively organized as the unit of analysis, most explicitly illustrated in his purposive-cognitive theory of learning.

Tolman's scientific contributions are contained in two books and approximately 100 articles published in scientific journals. His first articles, appearing in 1917–1918, while he was at Northwestern, were in the area of cognition, a foreshadowing of the strongly cognitive character of his later systematic formulations. These studies were, however, cast largely in the mold of classical introspective and associationistic problems—imageless thought, retroactive inhibition, association-time for pleasant and unpleasant words. As he later commented, "the behaviorist point of view had not yet really got into my blood."

It was soon after arriving in Berkeley that his thinking, stimulated partly by Watson's work, turned strongly toward behaviorism, but it was behaviorism with a difference. In 1922, Tolman's paper, "A new formula for behaviorism," argued for a "true nonphysiological behaviorism," one whose concepts would be appropriate to the psychological level of description and prediction and capable of bringing into a single framework *all* of the phenomena of psychology, including the subjective. He then plunged into this task, writing a series of articles which attempted to show how such 'mentalistic' concepts as sensation, emotion and consciousness could be translated into new objective behavioral terms,

These ideas, greatly expanded and elaborated, were finally brought together in his major book, *Purposive Behavior in Animals and Men,* 1932. Here he introduced the concept of behavior as *molar,* as having properties of its own, to be identified and described irrespective of the underlying physiological processes. This he set against the narrow molecular definition of behavior, typified in the "muscle-twitchism" of Watson. Tolman asserted that the distinctive features of molar behavior are its purposiveness and its cognitiveness. A crucial scientific task as he saw it was to treat these concepts of purpose and cognition objectively, avoiding the "taint of teleology or of subjectivism" by dealing only with *observables,* both on the side of the stimulus-situation and on the side of resultant behavior. He was thus an early proponent of a kind of operationalism (though he did not then so term it), conceived of independently of P. W. Bridgman.

In this book, Tolman presented his first major effort at a comprehensive theoretical scheme in which there would be place for all the kinds of variables thought to be determinative of behavior and for their complex interrelationships. The system conceived of mental processes as functional variables intervening between stimuli, initiating physiological states, and the general heredity and past training of the organism, on the one hand, and final resulting responses, on the other. These "intervening variables" were conceptualized as various capacities of the organism and various purposive and cognitive determinants, *e.g.* demands, means-end-readinesses, discriminanda-expectations, manipulanda-expectations and sign-gestalt-expectations.

Written in Tolman's inimitable style, replete with his inventive terminology, proposing a highly original theoretical system and undertaking to show how a great range of data from experiments on animal learning and motivation could be organized within this scheme, it is obvious why this book came to have a profound and sustained impact on the field of psychology.

In a series of articles that followed the book, Tolman elaborated the concept of intervening variables. As he conceived it, the study of behavior should properly be broken down into two separate sets of specifications: (1) the relationships between independent variables (*e.g.* stimulus-patterns) and intervening variables (*e.g.* purposes and cognitions); and (2) the relationships between intervening variables and the dependent variable of resultant behavior. This conceptual model widely influenced psychological thinking and Clark Hull, for one, adopted it, though the sorts of intervening variables he chose were quite different from Tolman's. In his presidential address in 1937 before the American Psychological Association—"The determiners of behavior at a choice-point"—Tolman made use of the scheme of independent-intervening-dependent variables to summarize and interpret the evidence from a great variety of animal experiments.

Tolman was a rare combination of system-builder and experimenter. Over the years he and his students carried out an active program of laboratory studies, mostly of rats. These studies may roughly be grouped under five main headings: (1) latent learning; (2) vicarious trial-and-error; (3) reward-expectancy; (4) 'hypotheses' in rats; and (5) spatial orientation and place-learning. All of these experiments were intended to be supportive of his theories, especially his central notion of the essentially cognitive character of the learning process.

In Tolman's view the important kind of learning is *sign learning*, the

learning of sign-significate relations. Learning is the acquiring of means-end-readinesses—beliefs about "what leads to what." The organism thus typically learns a cognitive 'map,' not a mere movement-pattern. Habits for Tolman were of the nature of sign-gestalt-expectations. As he and Egon Brunswik pointed out in "The organism and the causal texture of the environment" (1935), such expectancies, being necessarily based on probability-characteristics of the environment, are always provisional, that is to say, they are 'hypotheses.'

Basic to Tolman's thinking was the important distinction between learning and performance, a distinction most sharply pointed up by the influential and controversy-arousing experiments on latent learning in rats. Motivation he conceived of as affecting both performance and learning, but in quite different ways. Motivation drives the organism to performance, guided by existing cognitive structure. Motivation also facilitates the initial acquiring of the cognitive structure, but not through the simple S–R reinforcement resulting from need-reduction as asserted in the Law of Effect, but through the emphasis given by rewards and punishments, which aids the formation of sign-significate connections.

As a learning theorist Tolman primarily concerned himself with developing a comprehensive account of all the kinds of factors which govern learning; the details of the learning-process he left fairly vague. He rejected the notion of any form of monolithic learning-theory, such as S–R reinforcement or contiguity theory, maintaining, as for example in his 1949 article, that "there is more than one kind of learning." Here as everywhere in Tolman's systematic thinking there is evidence of his open-mindedness to the ideas of other psychologists and his use of them, no matter what the psychologists' theoretical persuasion. This catholicity was, however, no mere eclectic borrowing; in incorporating these diverse ideas, he transmuted them and fixed his distinctive stamp on them.

In studying the successive stages of development of Tolman's system it is easy to detect many of the sources of ideas that he sought to encompass. Indeed, he has himself pointed out some of them: Watson's behaviorism, the Gestalt psychologists, with their emphasis upon patterned wholes, Kurt Lewin and his concepts of valance, life space, etc., Egon Brunswik, with his stress upon the achievemental character of behavior. He was particularly receptive to ideas from his colleagues and students. Yet it was often he who in the first place stimulated in them the ideas for which he later gave them public credit. Though his system was unmistakably *his* system, he suffered no proprietarial paranoia about it.

He acknowledged also his debt to contributions from Freudian theory,

from personality theory, from clinical and social psychology, and from anthropology and sociology. Tolman always regarded psychology as both a biological and a social science.

His preoccupations with constructing a comprehensive scheme that could be extended to embrace all psychological phenomena, including the complexities of human thought and motivation, coupled with his basic concerns regarding various problems of man and society, led him in a number of writings to theorize about some of these broader psychological questions. In an article published in 1938 he stated the provocative view that just as an adequate psychology is a precondition for a complete physiology, so is an adequate sociology a precondition for a complete psychology. He sought in several articles to analyze the processes through which man's primary biological needs become transformed into secondary social needs and thus influence his behavior and the nature of the society in which he lives. In *Drives toward War,* published early in World War II, he took concepts derived from his thinking about rat behavior, combined with certain Freudian notions, and attempted to apply them to an understanding of the motivations which lead to war.

Tolman never ceased revising his system. His latest statement of it, full as always of innovation, appeared in 1959. The detailed architecture of the theoretical edifice changed in many features over the years, for it was always intended to be a highly functional design, responsive to new psychological problems and new findings; but the basic logic and plan of the edifice, with its twin foundations of purpose and cognition, remained firmly set throughout.

The fertility of the system can best be judged in terms of the rich variety of experiments to which it led, both in support and in rebuttal of his theories. Because the specification of relationships in the system was far from rigorous, these various experiments could not be regarded as crucial tests of the validity of the system. Tolman's intervening variables were more in the nature of qualitative generalizations that serve to categorize and sum up various empirically discovered relationships. He made few, if any, precise quantitative predictions from his theories. Indeed, he was skeptical in our present stage of psychological science of the value of mathematization, of refinement of measurement, of axiomatic and deductive methods. The true significance of his system lay, then, not in its power to yield logical deductions, but in its power to generate new ideas, to add new experimental questions, to call attention to unexplored regions of functional relationships in the comprehensive map of psychological processes that Tolman endeavored to draw.

Among the most significant contributions to psychology from Tolman's systematic thinking would seem to have been these: his conception of molar behavior and its essential purposive and cognitive character; his central use of the notion of expectancy; his view of the learning process as the acquiring of cognitive maps; his crucial distinction between learning and performance; his development and use of the concept of intervening variables; his insistence upon the necessity for encompassing all the different kinds of psychological variables in a single unified scheme. The pervasive influence of these basic ideas throughout contemporary psychological theory is plain to see.

Tolman the scientist was not separable from Tolman the teacher and the citizen. Throughout the many years in Berkeley his constructive influence was powerfully felt both within the University and within the community. He successfully led the protracted fight of the Berkeley faculty against the imposition of a loyalty oath. It is a fitting tribute to the dignity and integrity with which Tolman waged this struggle that in 1959 the Regents of the University of California conferred upon him the honorary LL.D. degree.

In the affairs of the Department of Psychology his wisdom, tact, and unfailing generosity contributed immeasurably to the well-being of the department. For his colleagues he was a constant source of personal strength and intellectual stimulation. In all these activities his wife, Kathleen Drew Tolman, played an indispensable supportive role for him. Theirs was an extraordinarily happy marriage.

Tolman's greatness as a teacher lay partly in the fact that he did not undertake to teach. It was not in his temperament to give a highly organized lecture presenting well thought through material; as he talked, new ideas continuously flooded his mind, new difficulties in what he was saying occurred to him, and he felt impelled to stop and worry out these problems and ideas right there and then. Students did not leave his classes with tidy lecture notes, but they did carry away with them enlarged minds and a vision of scientific psychology as an exciting intellectual enterprise.

Perhaps most of all they benefited from intimate contact with a creative mind. For Tolman indisputably displayed the essential traits of the truly creative person. His basic spontaneity showed itself not only in his cognitive processes but also in his general behavior. He was receptive to the world of ideas around him. He was sensitive to the subtleties of phenomena (perhaps the reason that Wolfgang Köhler once called him a "cryptophenomenologist.") Spatial imagery and metaphor (*e.g.* the "schematic sowbug") were for him indispensable parts of his creative thinking. He

sought to achieve systematic order but not at cost of undue simplification; he was able to tolerate a good deal of ambiguity and even disorder in seeking the larger, comprehensive unity.

Throughout his long career he showed a sustained commitment to his creative aims. He deeply valued the professional recognition of his work. Though he could view his work with detached humor, he never wavered in his deep conviction of the essential worth and validity of the undertaking. He did not find the act of creation easy. He had his proper share of creative misery; but despite this he found intrinsic joy in the enterprise.

Tolman's last published article, an address delivered by him in receiving the award of the American Psychological Association for "Distinguished Scientific Contribution," clearly exhibits the same freshness and flexibility of outlook, wry self-depreciation, and pleasurable anticipation of work to come that characterized all of his writing. Some quotations from the final passages of this last writing are typically Tolman:

> This is really all I have to say. It is not too brilliant an account; but I do want to point out that such experiments were fun to do, although they took a long time and although the results when we got them persisted in being slight, confused, and somewhat sleazy. They did give us, anyway, a beautiful chance to speculate about vector models, and this, too, was fun. But, whether such experiments or such models will in the end have any world-shattering importance seems doubtful. . . . But as it is, I am stuck with these sorts of data and these sorts of models, and I intend to go on playing with them. . . . In short, we will have a delightful time and absolutely no dull moments.

University of California, Berkeley RICHARD S. CRUTCHFIELD

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Psychology: A Study of a Science: Study 1. Conceptual and Systematic. Vol. 2. *General Systematic Formulations, Learning, and Special Processes.* Edited by SIGMUND KOCH. New York, McGraw-Hill Book Company, Inc., 1959. Pp. x, 706. $10.00.

This is the second of seven volumes projected by the Policy and Planning Board of the American Psychological Association with the purpose of analyzing the "methodological, theoretical, and empirical status of psychological science." As the title suggests, the twelve chapters by no means parallel one another. The "general systematic formulations" include essays by Dorwin Cartwright ("Lewinian Theory as a Contemporary Systematic Framework"), by Edward C. Tolman ("Principles of Purposive Behavior"), by Edwin R. Guthrie ("Association by Contiguity"), by Neal E. Miller ("Liberalization of Basic S–R Concepts"), by Frank A. Logan ("The Hull-Spence Approach") and by B. F. Skinner ("A Case History in Scientific Method"). In contrast to the broad scope of these contributions are three essays on particular approaches to or aspects of learning, "The Statistical Approach to Learning Theory" by W. K. Estes, "Learning Set and Error Factor Theory" by Harry F. Harlow, and "Rote Learning" by Arthur L. Irion. And in some other category or categories altogether are the last three chapters, "Some Recent Trends in Ethology" by R. A. Hinde, "Information Theory" by F. C. Frick, and "Linear Frequency Theory as Behavior Theory" by Douglas E. Elson. The diversity of these twelve offerings makes it impossible to review the volume as a whole by comparing the contributions in any way. It is true that the authors were given an outline with suggested themes of analysis, such as "structure of the system as thus far developed," "systematic and empirical definitions of independent, intervening, and dependent variables," "degree of programmaticity." These themes are, however, manifestly unsuited to the subject matter of some of the essays and the authors were permitted to follow them or disregard them as they chose.

The individual contributions must, in the final analysis, be evaluated by themselves, for they are unique. This will be done by readers of the book, for it can be pretty safely predicted that they will sample it piece-meal, rather than read it through. But a lone reviewer would be brash, indeed, to attempt such a program, for some of the offerings are highly specialized. I will attempt, then, to find a few general trends in the book as a whole. What is the science like, to judge from this book? How has it changed, from earlier days?

Murchison said, in his preface to the *Psychologies of 1925:* "We have here a genuine cross-section of contemporary theoretical psychology. Here are the norms with which future psychologies can be compared. Here are the principles that are up-to-date through the year 1925" (p. 1). The psychologies of 1925 abounded in systematic controversy. There were S–R theorists (Watson, Hunter), cognitive theorists (Bentley, Koffka, Köhler), and the advocates of motivation or purpose (Prince, and

MacDougall). These positions have their descendents in 1960, of course, but on the whole the systems have moved in the direction of Woodworth, who even in 1925 was fitting motivation and cognition into an *S–R* framework. In 1960, everybody incorporates motivation in one role or another (perhaps especially Lewin, Tolman, and Miller) and nobody wants to leave out thinking. Everybody wants to work with animals as well as people (except perhaps Lewin?), so the *S*s and *R*s are always in the picture.

This does not mean that there are no longer controversies. But the controversies are different. They seem to be mostly centered around "intervening variables" and methodological issues. Tolman, referring to intervening variables, says "This is where the schools differ" (p. 147). They may be many, as in Hull's elaborate system of constructs, or they may be absent, as in Skinner. They may be phenomenological ("expectations"), they may be pure constructs ("excitatory potential," "life space"), they may be "mediating" responses (either observable ones or "central responses"), they may be "sets," with status dubious. Just about all of the concepts which are new since 1925 fall in this area. They overlap—habit, inhibition, and drive are most frequently given some special title, definition, and systematic function. The diversity is still great.

The big issues, in 1960, appear to be methodological. The persistent questions are concerned with the right kind of theory-construction, experimental design, quantification, and statistics. An intense and self-conscious interest in construction of theory characterizes the book—not so much the substance of the theory but how it should be made. The "theoretical model" is a recurrent phrase (though Tolman calls it a "sloppy synonym"). Hull's logico-deductive method was the beginning of this trend, but except for Spence he has had few followers in this particular method. There seems even to be a trend away from logical formality. Miller's "liberalization" of the S–R approach is well-named, for he has departed widely from Hull in formal aspects. Lewin wanted a system of constructs, logically related, but he was even more anxious to treat "the full empirical reality of human experience" (p. 11), and thought violence was often done by the search for logical formulae and scaled variables. Skinner is an anti-formalist and the hypothetico-deductive method comes in for a good deal of sarcasm from him.

The really new chapters in the book, in the sense that they were utterly unpredictable in 1925, are the ones concerned with statistical models, information-theory, and linear frequency-theory. These are not theories in the old sense, but models geared for quantification. Quantification has come a long way since 1925, but even here one finds areas of disagreement. Hull, of course, was most anxious for quantification of his system; but still, many of the derivations had to be in terms of greater than or less than, instead of precise quantities. Tolman, as well as Lewin, has always been reasonably content with the greater-than language, and finds psychology today carried away into a flight into too much statistics and too great a mathematization. Skinner finds the cumulative curve, which makes behavior 'visible,' the ideal method of quantification, making statistics unnecessary. Miller believes worth while a "rigorous though qualitative approach which first determines the general significance of the variables, and then goes on to narrow down the choice toward a more exact quantitative function" (p. 286). Estes, in a 111 page chapter on the statistical approach to learning theory, is representative of a sizeable vanguard who

want precise quantitative description and empirical "moorings" defined in terms of response-probability. Despite the empirical emphasis, this approach requires a number of basic assumptions, such as random fluctuation of stimulus-elements. Quantitative predictions yielded so far apply only to experimental situations simplified enough to permit unambiguous identification of stimulus and response-variables. The novelty of the approach is not its new psychological insights but that "all linkages among constructs are explicit and determinate and all derivations of theorems are accomplished by an exact mathematical reasoning" (p. 472).

The chapter on rote learning seems a bit of an anachronism in this volume, since its author states that neither the problems nor the methods in this field have changed materially since Ebbinghaus' initial work. He sees hope for this area in greater standardization of procedures. But actually the new methods have penetrated to some extent—Hull's mathematico-deductive theory, statistical learning theory, and a few applications of information theory might be cited.

The chapter on ethology is also in considerable contrast to the theory-building and quantitative trends of 1960. The ethologist's aim is qualitative description and classification with quantitative work envisioned as a final and remote stage. Constructs are used by ethologists (*e.g.* "innate releasing mechanism") but there is little attempt at formal systematizing, just as there is rarely an atempt at precise quantification. Principles, when stated at all, are informal and are apt to be very close to the phenomena observed.

Contemplating these chapters together, it would seem that the major division in psychology today is between the theorists who seek laws within an elaborate system of constructs and those who seek laws of behavior which are above-all empirical. This is not to say that the former do not want data or that the latter have no constructs, but there is a great difference in emphasis between deriving a law from a group of carefully stated assumptions and finding it in the data.

How are these positions to be evaluated? The authors themselves differ markedly in criteria for evaluation. The criterion is "fun" according to Tolman, "knowledge" according to Guthrie, "control" according to Skinner, and fruitfulness in producing research or generating hypotheses according to others. Since this is only one of seven volumes, the appraisal of the science of psychology today can be left, happily, to some future reviewer of the whole.

Cornell University ELEANOR J. GIBSON

The Neuropsychology of Lashley. Selected Papers of K. S. Lashley. Edited by FRANK A. BEACH, DONALD O. HEBB, CLIFFORD T. MORGAN and HENRY W. NISSEN. *Introduction: Lashley and Cortical Integration* by EDWIN G. BORING; *A Salute from Neurologists* by STANLEY COBB. New York, McGraw-Hill Book Company, Inc., 1960, Pp. xx, 564. $9.50.

This memorial collection of Lashley's writings includes 31 articles selected from a total bibliography of 109 titles, not counting book reviews. Nearly all the articles are reprinted without substantial deletions. Students of brain and behavior who have had to consult dozens of different journals in the past will therefore find this volume exceedingly useful. If the 1929 monograph *Brain Mechanisms and Intelligence* is excluded as being widely available, this collection represents roughly one-third of Lashley's output. Articles were selected to "reflect Lashley's interests

and achievements, to avoid duplication and to hold the size of the final product within reasonable limits."

The editors, four distinguished students and associates of Lashley, express the hope in their Foreword that "this volume will stand as a modest tribute to Karl S. Lashley, man and scientist, and that it will constitute a useful addition to the scientific literature." A dual introduction, *Lashley and Cortical Integration*, by Edwin G. Boring, and *A Salute from Neurologists*, by Stanley Cobb, makes explicit recognition of Lashley's contributions, both to psychological and neurological thought. Professor Boring's paper is more than an appreciation. It traces the broad outlines of Lashley's professional career, cites the salient features in the development of his thinking and seeks to evaluate his impact on psychology. (Other memoirs of Lashley include that by F. Walshe in *Neurology*, 1958; Leonard Carmichael in the *American Philosophical Society Year Book* for 1958; D. O. Hebb in this JOURNAL, 1959; and F. C. Bartlett in *Biographical Memoirs of Fellows of the Royal Society*, 1960.)

Lashley's papers are arranged in chronological order with no attempt to re-group or categorize. The volume begins with *Notes on the Nesting Activity of the Noddy and Sooty Terns* (1915), a paper which demonstrated quite early Lashley's excellence as observer and recorder of protocols. Samples of early papers on inheritance in protozoa (1915) and salivary conditioning in man (1916) are given. The ablation and deprivation studies (1917-1943) are thoroughly represented, as are the critiques of the reflex and connectionism in general (1930). Experimental studies of sensory representation are covered (1934), as well as the somewhat cumbersome study of non-continuity in learning (1942), an issue which Lashley casually raised in the 1929 monograph. The last four papers represent Lashley's grand summary of thirty years of his experimental work and that of his students. The position of contemporary neuropsychology is evaluated and a masterful exposition of current problems is presented: the evolution of mind (1949), search for the engram (1950), serial order in behavior (1951), and finally cerebral organization and behavior (1958). The volume ends with a complete list of Lashley's publications.

A number of errors have crept into the text, but these are minor and do not detract from the value of the book. The chronological order of titles in the bibliography is occasionally violated and at least one title is actually omitted (Introduction to Paul H. Schiller's book on *Instinctive Behavior*, 1957). Date of publication of "In Search of the Engram" (1950), though cited in the bibliography and by Professor Boring in the Introduction, is omitted in the Editor's Note.

The danger in assembling a selected sample of articles for republication in book form is that the omissions, since they remain scattered, tend to become neglected. It is important then to inquire whether any prominent papers have been omitted. Exclusion of the 1929 monograph and the outstanding critique of cortical cytoarchitecture studies with George Clark is certainly justified by their great length. Unfortunately, the two short papers in which the reduplicated trace and interference pattern theories are developed (1931, 1942) are likewise omitted. Studies which might have found a place in the volume are the gold foil experiment with K. L. Chow and J. Semmes (1951), the anatomic contribution of thalamo-cortical relations in the rat (1941), and the effects of destroying the visual association areas in monkeys (1948). This latter, together with the earlier rat and monkey work, developed the model for subsequent ablation studies using Lashley's tests and brain

reconstruction techniques. The foray into the theory of libido (1924), though not an especially creditable performance, might have served to illustrate the wide range of Lashley's interests.

Though there could be no continuity between successive items in this volume, a number of threads stand out and can be followed in reading the papers chronologically.

(1) Sensory and motor factors in recognition. The breadth of Lashley's interests in this problem ranges from nest recognition by birds, the recognition of geometric pattern by rats reared normally and by rats surgically or environmentally deprived, to equivalence and transfer in the sensory and motor spheres throughout the phylogenetic scale. Klüver's deft handling of stimulus equivalence (1933) follows directly from his association with Lashley.

(2) Attention and the central autonomous process. Stemming from the critique of the reflex and masterfully developed in the Hixon Symposium paper delivered in 1948, the concept received systematic status in Hebb's book, *Organization of Behavior* (1949), and in Lashley's Vanuxem Lectures at Princeton University (1952), which were never published.

(3) Dynamic and integrative activity of the cerebral cortex. "The pattern and not the location of energy . . . determines its functional effect. Spatial and temporal order . . . appear to be almost completely interchangeable in cerebral action" (1951). Lashley was never satisfied with Köhler's rejection of connectionism, though he dismissed the idea of synaptic resistance (1924). A careful reading of the paper in the symposium on "Visual Mechanisms" (1942; not included in this volume) shows Lashley toying with an associative mechanism to achieve reduplication of the memory trace.

The anatomic studies of Lorente de Nó (1934) have revealed a system of cross connections in the cortex which will permit the spread of excitation in any direction along the surface. Many adjacent neurons are capable of mutual excitation and the whole system is organized as a network, with loops of various lengths and complexity, capable of transmitting impulses from cell to cell across the cortex, or of re-exciting initial points of stimulation by the action of return circuits having diverse characteristics.

"*From such a structural organization* functional properties may be inferred with some confidence. Excitation started at any point must spread from that point throughout the system. . . . If the system is uniform throughout, a series of radiating waves should be produced. . . . The timing of the waves should be uniform, *since it is dependent upon the speed of conduction and the refractory periods of the elements of the system.* With several or many points of excitation, interference patterns will be formed" (p. 312-313; italics mine). This is a field theory to be sure, but not one that depends upon the existence of electrochemical gradients in cortical tissue. Hebb's treatment of Köhler's field theory and equipotentiality together (1949) has tended to obscure this point.

Lashley paid only perfunctory attention to criticisms of his theoretical position. Though Hunter's objections to the concept of mass action did evoke a brief reply (1931; not included in this volume) and Spence's counter position of continuity was criticized (1942), the demonstrated limitations of equipotentiality which appeared during the 1950s never disturbed him to the point of meriting treatment in the

literature. On the other hand, Lashley wryly repeated to his students and associates that he himself had disproved all neuropsychological theories including his own.

Lashley's flair for the dramatic provoked much attention, controversy, and even humor, sometimes with mischievous intent to startle. He illustrates his materialistic philosophy with the following pithy statement entitled *The Will* (1958):

"Voluntary action is often cited as purely mental, chiefly because freedom of the will is indispensable to many established ethical systems. Voluntary action is usually defined as a choice between two foreseen alternative actions. This is beautifully illustrated by the spinal frog. If a bit of acid is placed on the midline of the rump, the preparation responds by alternate jerking of the hind legs. . . . Finally, the motor system of one leg dominates, and that leg performs the complete wiping reflex, while the other is extended. Here is foresight; we often see a man making such tentative actions while he is reaching a decision, and his verbal debating (foresight) about it is only a substitute for other tentative actions. After a period of indecisions, of vacillation, the spinal frog also reaches its decision. Save in the complexity of the alternatives, the procedures are the same."

Michael Reese Hospital
Chicago, Illinois
JACK ORBACH

Behavior Genetics. By JOHN L. FULLER and W. ROBERT THOMPSON. New York, John Wiley & Sons, Inc., 1960. Pp. ix, 396. $8.50.

"Our conclusion . . . is that we have no real evidence for the inheritance of traits." In 1924 Watson could honestly assert that. This timely volume chronicles how much has since been learned. "The . . . important generalization to be drawn from the variety of studies summarized . . . is the ubiquity of genetic effects. Quantitative measurement of behavior combined with genetic techniques finds an influence of heredity in insects, rats, and men, both in behavior which is relatively unmodifiable . . . and in behavior which is the outcome of a long period of development . . ." (p. 137).

The field of behavior genetics studies the genetic mechanisms underlying individual differences in behavior. It involves the measurement and analysis of behavioral differences and of the genotypic diversity among all members of biparental species.

The purpose of this pioneer work is to provide a "modern statement" of behavior genetics as well as a "comprehensive treatment" of its literature. The presentation is organized around three questions. "What are the effects of heredity on behavior? How large are these effects? What mechanisms are involved?"

The book results from the laudable collaboration of two men. Fuller, a biologist who went to Bar Harbor to study the physiology of behavior; and Thompson, a psychologist from Wesleyan University. The many excellent features will be considered first since they far outweigh the shortcomings.

The work's major contribution lies in a concretely objective account of the principal facts. The areas reviewed include sensation and perception, animal behavior, learning and behavior modification, intelligence, personality, temperament, and mental disorders. A large variety of problems in each area is soberly discussed and critically evaluated with many excellent methodological appraisals and suggestions. The supporting bibliography totals about 800 references. Their book is divided into

three sections (1) an exposition of the principles of genetics and the methods of behavior genetics; (2) a truly scholarly, comprehensive and informative review of literature; and (3) suggestions about theory and method.

There is a rich selection of topics from genetics covering many of the basic concepts as well as introducing the reader to population genetics and physiological genetics. The enlightening discussion of heritability as a property of populations and not of traits clarifies a widely misunderstood concept.

The treatment of methodology in studies of man seems to our nonspecialist eye like a sound introduction to a complicated field. The advantages and limitations of pedigree studies, Weinberg's propositus, Dahlberg's later sib, Hogben's a priori and Haldane's a posteriori methods plus the knotty problem of correcting morbidity rates for age are clearly outlined. Also covered are correlational and twin methodology and R. B. Cattell's powerful multiple-variance model.

Other outstanding features include a valuable discussion of (1) age effects in the context of two well studied substantive problems, the taster phenomenon, and audiogenic seizures; (2) the question of scale transformations illustrated with original data; and (3) the non-congruence model, which gives an enlightened picture of the kind of relation it is reasonable to expect between genes and behavior. Finally, the authors are to be congratulated for abandoning the unfortunate term psychogenetics and adopting the descriptive one, behavior genetics.

There are several shortcomings, some of which should be corrected in later editions. The exposition would have benefited from treating as dimensions the phenotypic characteristics with respect to which individuals differ. Much awkwardness could have been avoided if, instead of referring to "trait-bearers" and "non-trait-bearers" (pp. 49, 70, 112, etc.), reference had been made to different modes or degrees of expression of a trait. While the language employed might appear adequate for describing the distinction between tasters and non-tasters, it is awkward when when to distinguish (1) right-handed from left-handed individuals; (2) dichromats from trichromats; or (3) the subtypes among dichromats.

The symbols *G* and *F* are inconsistently used to label the generations of selection. Confusion occurs when *F* replaces *G* on pages 71, 214, 264, and 265 because *F* is correctly used throughout for strain crosses. Discussion of the symbolism problem would have helped since usage in the journals is also inconsistent. Fig. 9—1 requires further explanation because the white sections of the bars remain a mystery. P. 67 misleadingly states "The average genetic correlation between . . . parent and . . . offspring is 0.5. . . . The genetic correlation between ordinary siblings is also 0.5." The *exact* genetic correlation between parent and offspring is 0.5 whereas the full sibling correlation ranges from 0.0 to 1.0 and is only 0.5 on the average.

The Hardy-Weinberg law was discovered by an English mathematician and a German physician, not by two geneticists. The report of the association between Mongolism and trisomism should be qualified because population surveys are yet to be made to ascertain the proportion of trisomics that are mongoloid. It is incorrectly reported that, unlike Tryon, Heron selected genotypically. Heron and Tryon both selected extreme phenotypes without mixing their bright and dull strains. Further, the reader cannot possibly appreciate the trend of the early generations in the Heron study because of the omission of the foundation population values. *Both* strains are *brighter* than the foundation population for three generations.

In prophetic and insightful style, pages 344 and 345 suggest "a new approach to the genetic part of behavior genetics . . . manipulation of genotypes instead of phenotypes. . . . Studies of heritability will be carried out in the framework of population genetics." In all honesty this discussion would appear less mysterious and prophetic if there had not been a surprising omission of appropriate references and explanation at this point (*e.g.* Tryon, 1958 in the bibliography). The student of behavior is not shown the techniques of "manipulating . . . chromosomes" in the mouse where $N = 20$. At the same time there is an unfortunate neglect of the extensive *Drosophila* literature on the subject dating back at least to Mather's 1943 paper. That literature deals directly with, and provides the only concrete examples of, their "new approach."

While the coverage of genetics is comprehensive, the structure of genetics, one of the most elegant in all science, is not thrown into full relief by the presentation. Proper articulation of this structure would have made clear why it is *not* surprising that most behaviors show hereditary variation.

One gets the impression that the authors dodged the most challenging aspect of their task by placing their theoretical schema at the end. Skillful use of it throughout to organize and interpret the large body of data reviewed might have made a truly outstanding contribution. Furthermore they fail to extend the implications of their survey. Since genotypic diversity and its consequent individual differences in behavior are as widespread as they have shown, what does this mean for whole areas of psychology in which the experimental analysis of behavior proceeds on the counterfactual assumption that all organisms are alike? Also, the evolutionary consequences of the intimate connection between behavioral and genetical variation are overlooked. Behavior often determines which individuals survive and reproduce.

This book, nevertheless, has much to offer and will be a standard for years to come. It is being adopted for our behavior genetics course and added to the graduate reading list. In the authors' words possibly the "most significant contribution of behavior genetics is its documentation of the fact that two individuals of superficially similar phenotype may be quite different genotypically and respond in completely different fashion when treated alike" (p. 38). Clearly, detailed information of this nature must henceforth become part of the armamentarium of every serious student of behavior.

University of Illinois JERRY HIRSCH

Counseling and Psychotherapy: Theory and Practice. By C. H. PATTERSON. New York, Harper and Brothers, 1959. Pp. xii, 320. $5.00.

This book has been written to fill a need for a text to be used in an advanced course in counseling and psychotherapy. The author believed that no volume before his own was sufficiently comprehensive and systematic in preparing the students for their work in this area. Patterson's approach is based upon phenomenological psychology and client-centered therapy. He vigorously advocates a systematic point of view and feels that the actual choice is between psychoanalysis and self-theory. The influence of Rogers is very apparent, and the author believes that the client-centered approach contains, in its simplest and most basic form, the necessary and sufficient conditions for psychotherapeutic personality change. Patterson rejects an eclectic approach to the teaching of therapy because he feels no instructor can present dif-

fering points of view with equal effectiveness. He believes that one's systematic bias should be made explicit. The emphasis of this book is upon the theory and philosophy of an approach rather than upon techniques.

Generally speaking the author is faithful to what has become known as the client-centered approach. In his discussion of controversial issues, such as transference and countertransference, diagnosis and evaluation, and depth psychology, he not only clearly states the client-centered position but also gives useful scholarly discussions, drawing upon extensive readings and knowledge within the fields of psychology and psychiatry as well as cultural anthropology, education, personnel work, and psychoanalysis. After each chapter he has from 20 to 100 references to pertinent books and journals which will serve as guides to the serious student of counseling and psychotherapy. Patterson's style is clear and concise in spite of the fact that he discusses some highly complex issues.

The author claims that his volume is comprehensive. From this point of view the reviewer believes that at least one part of the book should have been devoted to the client's experience of psychotherapy. One unique characteristic in all of Rogers' writings is his ability to relate theorizing and actual, transcribed client verbalizations from the counseling hour. Patterson's book is primarily oriented toward the therapist, how his values, needs, and cultural background affect the therapy relationship. I question the wisdom of moving away from an emphasis upon what is actually happening during the counseling hour. Another related criticism is that Patterson gives the impression that therapy draws more upon the therapist's cognitive than his emotional-personal qualities. The whole spirit of the book is discursive and rational. The reviewer would recommend that some of Rogers' writings be added to the present volume for advanced courses in counseling and psychotherapy.

Now some brief comments shall be made about each part of the book. In Part I, the author argues that the same basic therapeutic attitudes are applicable in all facets of personnel work. The techniques used, however, may vary whether the purpose at hand is vocational-educational counseling, or therapeutic counseling and psychotherapy.

Patterson argues that training in therapy should come at the earliest practicable moment. Rogers claims that the most powerful teacher is the student's own direct clinical experience. The reviewer gets the impression that the author advocates an intermediary position between cognitive and experiential learning. He also wants to help the students to develop a 'professional attitude' and a 'systematic, theoretically based approach.'

Part II is a discussion of the background for counseling and psychotherapy and includes chapters on ethics and counselor needs, values and psychotherapy, and cultural factors and psychotherapy. Apparently these ideas are close to the heart of the author. In the volume as a whole he presents a distilled and updated client-centered position, but in these chapters he offers stimulating discussions of his own.

It is pointed out that no therapist can be completely value-neutral. The author's ideal seems to be to leave the client as free as possible within the limits of certain basic values which are inherent in the therapist-client relationship. Rogers never focused much upon the client in his social-cultural context. Patterson, however,

finds that it is rather important that the therapist be sensitive to this cultural content of the client's communications.

Part III, the largest in the book, deals with significant aspects of the therapeutic relationship. First, he discusses in general terms the understanding and manipulative approaches to human relations. The understanding approach reflects a positive, optimistic, encouraging view, while the other is based on a negative, pessimistic, discouraging view. Second, he contrasts the eclectic and systematic viewpoints in therapy and continues to present the broad outline of the client-centered theory of therapy and personality.

The latter half of the book is concerned with the practical implementation of the client-centered position. He discusses such topics as counselor expectations, rapport, listening, understanding, and non-verbal communication. This reviewer found his discussion of non-verbal communication to be a real contribution to stating a comprehensive client-centered approach.

In regard to diagnosis and evaluation Patterson defends the client-centered position in a well-documented chapter. "Diagnosis or evaluation is not considered to be a necessary or essential condition for counseling and psychotherapy. Nor are tests considered to be useful or desirable parts of the therapeutic process. This is in contrast to vocational and educational counseling, where tests are useful in most cases." But even in education-vocational counseling it is thought to be possible to apply the same basic attitudinal orientation toward the client, even if tests are applied. Rather than prescribing tests for the client it is perhaps desirable to enlist his coöperation in selecting them.

Part IV, the last in this volume, is a discussion of three controversial questions. I found the first question "Is depth psychology necessary?" somewhat misleading. It is my guess that Rogers would not agree that client-centered therapy lacks depth. In a recent paper Rogers has even conceived of the essence of therapy as deep, experiential moments of personality integration. An alternative heading would be: "Is interpretive therapy necessary?" In relation to this question the client-centered answer would be "No." The two other questions which are discussed are "Psychotherapy: art or science?" and "Common elements in psychotherapy; essence or placebo?"

In conclusion this reviewer would like to say that he found Patterson's book both enjoyable and stimulating. On the whole, it advocates a fairly orthodox, but up-to-date and scholarly presentation of the client-centered approach to counseling and psychotherapy. It is a book written by a teacher of counseling for his students. In this sense it is student-centered rather than client-centered. The volume is somewhat lacking in giving the reader a feeling for what the *actual experience* of therapy is all about; it is more academically than clinically oriented. Patterson's contribution deserves, however, to be widely read and used for advanced courses in counseling and psychotherapy.

Cornell University LEIF J. BRAATEN

The Sage of Sex. A Life of Havelock Ellis. By ARTHUR CALDER-MARSHALL. New York, G. P. Putnam's Sons, 1960. Pp. 292. $5.00.

Mr. Calder-Marshall, free-lance writer and biographer, has written a sensitive,

scholarly book which will be a welcome addition to earlier biographies of Ellis as well as his own account, *My Life.* The author does not attempt to evaluate or describe the scientific and literary work of Havelock Ellis, but concentrates on the nonprofessional side of his character in an effort to understand the complex personality of the pioneer who so profoundly influenced the twentieth-century social attitudes toward sexual behavior.

As a young schoolmaster in Australia, Ellis was tremendously impressed by James Hinton's *Life in Nature,* a rather rambling and loosely reasoned attempt to reconcile scientific materialism with Christian religion. Hinton was a physician who recommended, among other social reforms, a degree of sexual freedom quite out of harmony with the mores of nineteenth century Britain.

Later in life Ellis referred to his "conversion" at the age of nineteen. This involved the influence of *Life in Nature,* but was also based upon an important decision which was to have far reaching consequences. Ellis came to the conviction that it was unnecessary to worry about whether one is this sort of person or that sort of person. The important thing is to accept oneself as one is, and then devaluate the self in the interest of serving others. The behavior in which people indulge is unimportant. The motivations underlying behavior are of basic significance. This philosophy guided his life, and greatly facilitated his work in the area of sexual psychology. At the time of the "conversion" Ellis chose the study of sex as his vocation, and to that end began his training as a doctor of medicine.

One of the most widely read men of his day, Ellis was not always a critical reader: he described his brain as functioning as a kangaroo jumps. He was, of course, a prolific writer whose contributions touched many fields in addition to his principal one. An abiding concern with social reform is reflected in several of his books including *The Criminal* (1890), *The Nationalisation of Health* (1892), and *Marriage To-day and To-morrow* (1929). Other writings deal with studies of genius, race degeneration, the interpretation of dreams, Spanish sonnets and folk songs, and English drama from Marlowe to Shaw.

One of the earliest of his books on sex was *Man and Woman.* First published in 1894, this classic went through six editions, the last appearing in 1926. Because of the exceedingly advanced, not to say socially premature, ideas they contained, Ellis' writings inevitably aroused opposition and criticism. His book, *Sexual Inversion,* was rejected by English publishers—partly, in all probability, because Oscar Wilde's trial and conviction were still in the forefront of the public mind. The volume was promptly published in Germany. Later an edition was published in England, but the publisher was arrested, and Ellis decided against publishing any of his *Studies in the Psychology of Sex* in Britain.

As a result of his writings, the liberality of his views with respect to extramarital sexual relations, his tolerance toward inversion and other "anomalies," etc., Ellis' contemporaries generally took it for granted that his personal life was free of the usual amount of frustration and inhibition. The fact was that until he was sixty years old Ellis was uncertain of his sexual potency, and in his liaisons with various women he stressed the importance of "affectionate intimacy without passion."

In 1897 he entered into correspondence with a woman in South Africa who had written a book which reminded him pleasantly of his Australian sojourn. After

several months Olive Schreiner came to England to meet Ellis. She met a handsome, exceedingly shy young man with a squeaky voice and no sexual experience. Miss Schreiner was not lacking in responsiveness nor in experience, but although the association lasted for two years, and resulted in the development of deep, mutual affection, physical intimacies never progressed to the complete consummation she desired.

Ellis met Edith Lees in 1889, and she found him rather unattractive. Later she read *The New Spirit* (1890) and was strongly impressed. Their next meeting was the first of many, and Ellis reached the conclusion that with this woman he could establish a successful marriage. The meaning of "successful" he explained at the time of his proposal, and it was, for that day or this, quite novel. (1) There were to be no children. (2) Neither individual was to be financially dependent upon the other. (3) Each partner would maintain a separate domicile. (4) There was to be no element of possessiveness, and jealousy was foresworn. The matter of physical intimacy seems to have been left open for future developments, but the positive value of marriage was to be spiritual companionship, and Ellis added that he had no passionate feelings.

Edith accepted the conditions, paid her half of the cost of the wedding ring, and the couple went their separate ways immediately after the civil ceremony. That evening Edith held an informal reception in her flat, but Ellis sent his regrets and only reluctantly agreed to attend after all of the other guests had left.

Ellis felt that frequent separations contributed to the permanence of his marriage, and he encouraged Edith to take vacations abroad. On one such occasion she wrote of meeting a feminine friend for whom she found she felt a passionate attachment. True to his philosophy, Ellis told his wife she should accept herself as she was (his diagnosis was congenital inversion), and act as she felt she should act. With his "blessing," which was probably not what she really wanted, Edith entered into the first of several homosexual liaisons which occurred sporadically throughout her life.

For his part Ellis had several female friends at different times during the twenty-six years of his marriage, but although there was strong attachment and mutual respect, the relations seem to have been confined to mildly erotic exchanges marked by affection rather than extreme passion. All of these "affairs" were dutifully reported to Edith, who, if not enthusiastic, was at least determined to remain Ellis' wife.

Edith died in 1916 of "cyclothymic depression" and diabetes. Ellis cared for her as best he could while she alternated from suicidal depression to periods of euphoria. During the latter, she contracted heavy debts while embarking on one grandiose scheme after another. In connection with one of these projects Edith engaged a young French woman, Françoise Cyon, to translate certain English works into French.

Because of their peculiar marriage contract, Ellis was not legally liable for any of the debts incurred by his wife during her mental illness, but he was determined to pay what he could in any case where serious hardship was involved. To this end he corresponded with every creditor, and it was revealed that Madame Cyon was destitute. Ellis wasn't much better off, but after several months of correspondence

Françoise came to England and found work of sorts. Ellis attempted to help her, and although he was twice her age she gradually fell deeply in love with him and declared her affection.

Ellis preferred a nonphysical relationship, explaining that he would not make a satisfactory lover. He also urged reconciliation with her Russian husband who had not been contributing to her support. The reconciliation failed, and the sixty year old man and thirty year old woman became lovers in fact. Ellis was astonished to discover his own sexual normality, and credited it to Françoise's unselfish love and devotion. There followed more than a quarter of a century of perfect, monogamous "marriage," which endured until Ellis' death in 1939.

From this carefully drawn portrait Havelock Ellis emerges as a fertile though not incisive thinker, a person dedicated to the improvement of human relations, an unselfish but fiercely independent individual, a man who, though limited in self-understanding, possessed a great capacity for the sympathetic understanding of others and their problems, Above all, perhaps, Ellis is portrayed as a lovable, often eccentric human being possessed a great compassion, tolerance, and love of his fellow men.

University of California, Berkeley FRANK A. BEACH

The Mother-Child Interaction in Psychosomatic Disorders. By ANN M. GARNER and CHARLES WENAR with JENNIE B. KAHN and JEAN P. CHAPMAN. Urbana, University of Illinois Press, 1959. Pp. viii, 290. $6.00.

"Fortunately it is not necessary to know everything about everything in science," declare Garner and Wenar, "before one can learn something about something" (p. 174). As one reads through the details of the research project reported in this volume, one is likely to conclude that the senior authors and their collaborators have learned something about mother-child relationships and that they have learned almost everything about the problems and frustrations of clinical research.

The central hypothesis of the research, derived from psychoanalytic theory, was that "susceptibility to psychosomatic illness in children develops in the first year of life, when somatic response patterns are first laid down" (p. 11). This susceptibility develops within the context of a specific type of mother-child relationship in which the mother is lacking in what the investigators term "motherliness," a characteristic which is present to the degree that both the infant and the mother derive gratification from the processes of infant care. The research design involved the comparison of psychosomatic, neurotic, and physically ill groups, each composed of 26 pairs of mothers and their children ages six to twelve years. Assessment techniques included several brief indices of intelligence, both objective and projective personality assessment techniques, brief observations of the mothers with their children, and extensive mother-interviews. Standardized tests were used when possible, but several novel techniques were devised for the study.

Considering the complexity of the problem and the difficulties invariably attendant to such research, it is not surprising that only some of the specific predictions were confirmed, although there were few completely negative findings. Most readers will probably find the authors' interpretations of the results well within the bounds of scientific propriety. Mothers of psychosomatic children are epitomized as "ambitious, controlling women who have high expectations for their child

during pregnancy but find the actual caretaking of the infant unrewarding or disagreeable; because of their emotional investment . . . they are irresistably drawn to this ungratifying activity to the point of becoming entangled in a close, mutually frustrating relationship" (p. 160). In contrast, mothers of neurotic children are "relatively free in their approach to infant care and quick to express feeling, either positive or negative. . . . In the face of the highly charged negative relationship which develops both mother and child erect protective defenses to distance themselves from one another" (p. 165). The mothers of the physically ill children "are the freest and most relaxed of the group. . . . The over-all picture is one of realistic, uncomplicated pregnancy anticipations fulfilled in positive ways by the events of the first year of the infant's life" (p. 167).

The authors have been thorough and candid in discussing the shortcomings of their study. They point out that they have attempted "to preserve the clinical significance of the material without sacrificing methodological and statistical considerations" but that at every stage "this has proved a difficult task" (p. 5). The problems were indeed numerous. It was difficult, for example, to derive practical working definitions of some of the concepts central to the study. The attrition rate of the subjects was uncomfortably high. Reliabilities of the measurement techniques were often low and the degree of agreement between judges was at times unsatisfactory. It was, at some points, extremely difficult to be sure that a particular technique was measuring the variable for which it was designed.

An important strength of the design was its comparison between psychosomatic, neurotic, and physically ill groups. One is, however, inevitably tempted to wish for even more information, *e.g.* for comparable data on a group selected at random from the general population. Interpretation and generalization of the results would be clearer had we some notion as to how deviant these patterns are from 'normal' mother-child relationships.

The authors' question, and probably rightly so, whether they have always "been perceptive and flexible, or whether they have been merely arbitrary" (p. 177). Their answer is beyond reproach. "All one can do," they conclude, "is make the facts public and trace the steps of integration clearly" (p. 177). This they have done admirably and the result is a book which is both stimulating and enlightening. Their work will, of course, be of interest to those who are involved with the specific problems of mother-child relationships in psychosomatic disorders, neurotic disorders, and physical illness in children. It can as strongly be recommended, moreover, to any psychologist who is interested in the conduct of clinical research.

University of North Carolina HALBERT B. ROBINSON

Group Dynamics: Principles and Applications. By HUBERT BONNER. New York, Ronald Press, 1959. Pp. viii, 531. $6.50.

The burgeoning field of small group research is neither integrated nor summarized by any recent textbook. Cartwright and Zander have edited an excellent collection of introductory essays and representative empirical studies in their *Group Dynamics,* second edition, 1960, and there is also a more advanced, but equally excellent set of readings edited by Hare, Borgotta, and Bales, entitled *Small Groups,* 1955. (Reviewed by D. Cartwright, this JOURNAL, 69, 1956, 502.) Many teachers, however, hope for an integrative textbook.

The book under review will not serve this need. It is intended as "a comprehensive treatment of the dynamics of small-group behavior . . . designed as a textbook for college students and for professional reading by psychologists, social scientists, teachers, and personnel managers" (p. iii). Unfortunately, the book will prove worse than useless to any of the foregoing groups.

College students will be confused by the frequently encountered vague or meaningless statements. Example: "When we speak of the group . . . we think of it, not as a generalized entity, as sociologists do when they refer to an institution, but as an interaction system, as a dynamic and organized totality" (p. 40). They will be misled by poor definitions. Example: "A region is a segment of social space, an element of a social field" (p. 42). No student would ever guess from this that regions refer to *activities* rather than to areas. They will be amused at value judgments which abound in a supposedly scientific treatment. Example: "Power is wicked and immoral only when it is abused—when it exploits, divides, and disorganizes" (p. 280). If at all discerning, they will be confused by numerous inconsistencies. Example: "Irrespective of an individual's psychological biography, he behaves in accordance with the properties of the group at a particular time" (p. 25) as compared with "Regardless of their contacts with different racial groups some individuals find in their hatred of them convenient channels for the expression of deeply rooted aggressions" (p. 138).

Psychologists and other social scientists will be disturbed by the virtual lack of methodological criticism. Case studies, field studies, experiments with good controls, experiments with no controls, observations by such popular writers as W. H. Whyte (*The Organization Man*), and clippings from the *Christian Century* or the *New York Times* are given equal evidential status. Sometimes crucial experimental details are lacking, as when we are told that the subjects were "required to learn the words as a group" (p. 122), but are not told whether that means learning to five perfect trials by all members, learning to one successful reproduction by a majority of members, or learning by the brightest or most articulate member. Elsewhere, in contrast, we are given useless details. Example: We are told that a study was done at the "Harwood Manufacturing Company, in Marion, Virginia" (p. 145). Much more serious are numerous examples of statements that are either absolutely false or else demanding of serious qualification. Example: "Unlearning attitudes, like learning them, can take place only when there is interaction or communication between members or between groups. Intergroup hostility, accordingly, may be best eliminated if the hostile groups are brought into contact with each other" (p. 97).

Teachers of courses in small groups will be especially startled by the poor coverage. To be sure, this is a tremendous area to review, but there is clearly no justification for the claim stated in the concluding chapter that "In the foregoing chapters we presented a detailed exposition and analysis of the most significant problems of group dynamics" (p. 481). The chapter on "Human Relations in Industry" may be referred to by way of illustration. Only the first of the numerous studies conducted in human relations by the Survey Research Center of the University of Michigan is mentioned; this, of course, means that Seashore's extremely appropriate monograph (*Group Cohesiveness in the Industrial Work Group*) is not cited. The chapter does not mention any of the Ohio State leadership studies; nor the famous Coch and French study of representative, participative, and directive

methods in work change; nor Bavelas' work on communication nets. Most surprising of all is the lack of any mention of unionism. Incidentally, although published in 1959 the book contains no references (in any chapter) to any items written in 1958, and only refers to four items written in 1957.

Finally, few practitioners will find this book to their liking. In addition to many of the foregoing points, they are likely to object to the frequent attacks on straw men. Examples: "We have been critical of the argument that leadership in a democracy must be completely shared" (p. 277). "Inconsistency of training leaders by persons who exalt leaderless groups uncritically" (p. 196). "The leader is often conceived as a towering individual who by virtue of his 'gifts' stands above the 'crowd' and exhorts, compels, or persuades it to do his bidding" (p. 482). Both the practitioner and the student may gain the impression that such assertions are common in the scientific group-dynamics literature, but this is simply not so.

In summary, this is not a good book.

Cornell University LEO MELTZER

Psychology of Adolescence. By LUELLA COLE (in association with IRMA NELSON HALL). Fifth edition. New York, Rinehart & Co., Inc., 1959. Pp. xviii, 731. $7.00.

This latest edition of one of the widely-known textbooks in its area deals with a very wide range of subject-matter, from the relatively simple and straightforward materials of physical development to the rather tenuous materials of religious beliefs and moral behavior. The book has much to recommend it. Its author, having had long experience in writing, knows how to communicate her ideas and her subject-matter. The volume is comprehensive. Dr. Cole has incorporated the findings of a great deal of research and cites her sources, although, with very few exceptions, she has omitted names from her index—a glaring deficiency. She wisely avoids excessive and unnecessary use of professional and technical terminology, especially in the discussion of 'maladjustment,' deviant behavior, and problems of personality. The chapters on physical development are of superior quality and are appropriately illustrated. Throughout the book, there are many tables, charts, and other illustrations; but this volume is certainly not a "tabloid-text." The illustrations are relevant to the text.

This book, however, has its weaknesses, but is there a textbook about which this could not be said? Some of the discussions are too cursory and unenlightening, especially those dealing with tests of intelligence and of personality, including projective techniques. These discussions give the student no psychological insights into the tests themselves or into their uses for obtaining an understanding of the behaving, the dynamic human being. Dr. Cole merely describes the tests in a somewhat perfunctory fashion, without attempting a critical evaluation of them in relation to understanding "the adolescent personality." At this point, it is pertinent to ask whether there should be, in this type of book, any sort of detailed presentation of tests of intelligence and of personality. We find the same subject-matter included and repeated in books on childhood, adolescence, personality, and general introductory psychology; and, also, in books on guidance, counseling, industrial and clinical psychology. Isn't the student in psychology going to become satiated by such duplication and repetition?

In this book Dr. Cole, whether wisely or unwisely, is little concerned with

theory; nor is she always concerned with problems of basic causal factors in behavior. The materials are, for the most part, presented in a descriptive and nontechnical way.

There are, regrettably, some unfortunate, or misleading or incorrect statements in the text: "There is nothing fixed about it [intelligence]. It goes up and down with circumstances" (p. 141). "It may be that native ability will be shown to grow steadily up until senescence" (p. 143). "This child [with an *IQ* of 122] has 22 percent more mental ability than one expects from a child of his age" (p. 146). "By the time they [imbeciles] are adults they can complete about the first two grades of school" (p. 238). "She [a teacher] may be called upon to administer tests of personality, or she may require for her own guidance to use such measures" (p. 304). "A copy of a given pupil's test [of personality] may be sent to five 'experts' for interpretation, with the resulting assembly of five more or less divergent opinions" (p. 312). On pages 320-321 two pictures from the Symonds *Picture Study Test* are reproduced; but each of these is labelled as a "TAT Picture," although attributed to Symonds.

Such questionable and erroneous statements and details do not necessarily disqualify a book; but they do suggest that there are some sections that should not be read independently by the naïve student. On the whole, however, this is one of the more useful and desirable textbooks in its field.

Cornell University FRANK S. FREEMAN

Psychology in Theory and Practice. By THOMAS A. RINGNESS, HERBERT J. KLAUSMEIER, and ARTHUR J. SINGER, JR. Boston, Houghton Mifflin Company, 1959. Pp. xii, 480. $6.00.

Most textbooks for the introductory course in psychology lack a direction or a point of view. This new book by Ringness, Klausmeier, and Singer does not; it may be added to that small group which do have a special emphasis. This book might be particularly useful to the instructor who wishes to stress social psychology and its related areas. For this reviewer, however, there is a lack of content which makes the book difficult to appraise favorably.

The authors state as their objective the presentation of "major principles and theories rather than an encyclopedic body of specific facts." They succeed quite well in reaching that objective, though a few more facts might have been helpful. The instructor who elects to use this text must therefore add the specifics he feels are desirable, either through lecture or supplementary reading materials. For the present reviewer, considerable expansion of the sections on sensation and perception, learning, physiological psychology (which is all but missing), and several others of the classical content areas would be desirable. Otherwise, he could not feel that his students were getting a fair picture of the what and why of modern psychology. Without those additions, students would undoubtedly find the book easy to read, for it is extremely well written, but would come away with the impression that work in this field is directed toward producing a scientific sound version of *How to Win Friends and Influence People.*

If there is a major fault to be found in this text, it is one so intimately tied to the philosophy of the book as to make its correction impossible: There is a tendency to generalize basic principles (to make them meaningful to and directly applicable by

the reader) to a point far beyond that to which many of the data involved will stretch. The over-generalization is not often flat-footedly stated by the authors, but the idea that all psychological research findings should be directly and immediately usable by the student for his own problems is implicit in most of the writing.

Particularly susceptible sections are those on "Transfer of Learning," "Improving Perception," "Improving Communication," and "Motives and Emotions in Learning." In these sections particularly, the emphasis is on the way the student may study better, or read better, or listen better, or grow up better. Too often, the impression is left that psychology's sole purpose is the improvement of the study habits of freshmen. Such an approach will most likely make students report that the course was interesting (one goal of the authors); it may even serve to recruit more students as majors in psychology, but it is not likely to prepare those it recruits for the shock of discovery that this field is not solely concerned with their learning to be successful students or businessmen or citizens. The approach is, perhaps, analogous to that of the course in methods of teaching arithmetic in which the prospective teacher learns the tricks involved in getting children to manipulate numbers in the desired fashion, but never learns any mathematics. Just as that approach has turned out a generation or two of people who do not (and would rather not) know or care what mathematics is about, so this non-content approach to psychology may turn out a group of people who know all about psychology except what it is and what its goals are.

The University of Texas — JERRY V. TOBIAS

Die Sprache der Zeichnung. By CARLOS J. BIEDMA and PEDRO G. ALFONSO. Bern, Switzerland, and Stuttgart, Huber, 1959. Pp. 110, plus 22 pages of reproduction of pencil sketches and a packet of materials for testing. DM 8.50.

This book should be of considerable interest to psychologists working with such tests as the Rorschach. It is so clearly and simply written that a reader with a limited familiarity with the German language may comprehend it. The testing method, originated by Wartegg, was slightly modified by Biedma and applied by him and d'Alfonso during their 17-yr. terms as teachers in various schools of the Argentine Republic to 2812 persons of age 3 yr. to middle-age. Chief emphasis was placed, however, upon *S*s of high school age as it would be in the United States.

The method consists in laying before the *S* tested 16 black-framed squares of white paper (4×4 in.), together with a middle-hard pencil; and asking *S* to *add to something already on the square* whatever he wishes, with the end in view that the outcome in each case be "a sort of a whole thing" pictured. The things already there vary in simplicity, beginning with a mere point and continuing to as many as four lines in various relative positions, and also one or two circle-segments in various positions. The results will often appear as very humorous to the reader. The authors have developed rules for arriving at a series of values in 10 grades, for the following character-features: Mental liveliness, analyzing, clarity, phantasy, objectivity, observance, organizing, synthesizing, emotionality, confidence, unreservedness, sensitivity, perseverance, self-control, decisiveness, initiative, energy, adaptiveness, selflessness, friendliness, cautiousness, patience, social responsibility.

It is a fine example of bookmaking on the publisher's part.

Miami, Florida — MAX F. MEYER

Developmental Psychology: An Introduction to the Study of Human Behavior. By FLORENCE L. GOODENOUGH and LEONA E. TYLER. Third edition. New York, Appleton-Century-Crofts, 1959. Pp. xix, 552. $6.00.

The first edition of this textbook, by Dr. Goodenough, was published in 1934; the second in 1945. For the revision of these, resulting in this third edition, "Dr. Tyler has been entirely responsible," states the original author; since, she further says, "almost complete loss of vision made it impossible for me to prepare the new work myself." Recent research has been included, Piaget's views have been given more prominence, psychoanalytic materials and Freudian-based theory have been included, and the adult years have been given increased emphasis, this last being an altogether reasonable thing to do in a book on developmental psychology. It is gratifying to find authors who regard adulthood as an integral part of human development. Perhaps "the child" and "the adolescent" have had more than "their day."

The beginning student in this branch of psychology, for whom the book is intended, will find the revision of this well-received textbook to be comprehensive, clear, and its viewpoints adequately buttressed by research findings, even though not all psychologists will agree with all of the theories held and interpretations made; but where does one find a book in psychology which will evoke agreement, without dissent or criticism from the author's fellow-psychologists?

This edition is a worthy successor to its predecessors written by a distinguished scientist.

Cornell University FRANK S. FREEMAN

Hypnosis: Fact and Fiction. By F. L. MARCUSE. Baltimore, Penguin Books, 1959. Pp. xv, 221. $0.95.

The author states that his purpose is simply to separate fiction from fact in the field of hypnosis. His book, an original edition in a paperback is comprehensive in scope and clear in style. Marcuse succeeds in his stated objective in so far as this is possible in a field where experimentation is notoriously difficult and authorities themselves are in rather sharp disagreement.

His disposition of fiction will make the book of special interest to the layman. While his consideration of the 'facts' is naturally colored by his own perspective, and while the individual psychologist may take issue with Marcuse's conclusions, the reader may find Marcuse's review of this field quite worth while.

Colgate University G. H. ESTABROOKS

Seelische Gezundheit. By WALTER BETTSCHART, HEINRICH MENG, and ERICH STERN (ed.). Bern, Hans Huber, 1959. Pp. 356. DM 34.

A collection, very comprehensive, of 32 essays on psycho-hygiene by 34 authors, 23 in German, 7 in French, 1 in Italian, 1 in English. Presented in 10 groups under the subtitles (in English translation): General problems in psycho-hygiene; Marriage and family; Children and youths; Therapy; Psychosomatics; Law courts and judges; Anthropologic relations; Advertising and propaganda; Shelter need of any animal; Society's responsibility and task.

The collection reminds us in one of its early pages that the human fool, after his senseless jump into misfortune, displays his genius in struggling out of it. That is

the trend of the reasoning of these authors. They are worth reading if the languages are no obstacle. In this respect a Swiss has usually the advantage over an American. To select one or a few of the essays as particularly recommendable is for the reviewer an impossibility. They are all good in content, interesting in their appeal to any intelligent reader, and written in good style. The book is excellently printed and bound.

Miami, Florida MAX F. MEYER

Introduction to Exceptional Children. Third edition. By HENRY J. BAKER. New York, The Macmillan Company, 1959. Pp. ix, 523. $6.50.

The third edition of this well-known textbook is intended to meet the interests and needs of educators, primarily, and of other professional groups in the study of exceptional children of all types, including those having sensory and physical handicaps, as well as the deviates in intelligence, in both directions from the central tendency. This volume presents pertinent materials from each of the several special fields concerned: psychology, social work, medicine, education. Like its predecessors, this new edition is on an elementary level throughout; and as its title suggests, it should serve as a useful introduction to this important educational and psychological specialty.

Cornell University FRANK S. FREEMAN

Child Psychology. By ARTHUR T. JERSILD. Fifth edition. Englewood Cliffs, N.J., Prentice-Hall, Inc., 1960. Pp. xxi, 506. $9.65; text edition $7.25.

Jersild, in this edition of his widely-used textbook, has not only incorporated newer research findings, but he has placed much more emphasis, than in earlier editions, upon "ego psychology," upon interpersonal relations, and upon the "inner" world of the child as well as upon the "outer." He has also made more use of the contributions of psychoanalysis. Although the book contains some segmental treatment of psychological functions, which I believe is inevitable, Jersild has considerably extended and intensified his attention to the dynamic character of human development and to the significance of a wide variety of cultural forces.

Perhaps my major criticism of this text is that many of the topics and problems are dealt with so briefly as to be over-simplified and inadequate for professional students of the subject—even beginners. Here again, however, this is an inescapable result if an author attempts to 'cover' his subject within the space of an ordinary textbook, even by contemporary standards of size. The only other choices are to expand the book beyond the size acceptable to publisher and prospective instructors; or to limit the number of topics dealt with and present these quite thoroughly. Within the limits set by the pattern of his volume, Jersild has presented a readable, comprehensive, and informative treatment of child development and behavior.

Cornell University FRANK S. FREEMAN

The Importance of Wearing Clothes. By LAWRENCE LANGE. Hastings House Publishers, New York. 1959. Pp. 349. $7.50.

This amusing and interesting book has a wealth of examples and illustrations drawn from a wide variety of sources. The basic theoretical position follows that of Adler. Langer believes that the most important function of clothing is to demon-

state superiority both over animals and over other men. This he illustrates by examples from anthropology and from history of costume.

Langer's association with the theater; experience with patents relating to clothes, and the dress industry; his study with Veblem and Adler have given him many sources from which to draw. The book covers a wide variety of topics, including clothes and their relationship to sex, religion, government, communication, law, man's behavior, and the performing arts.

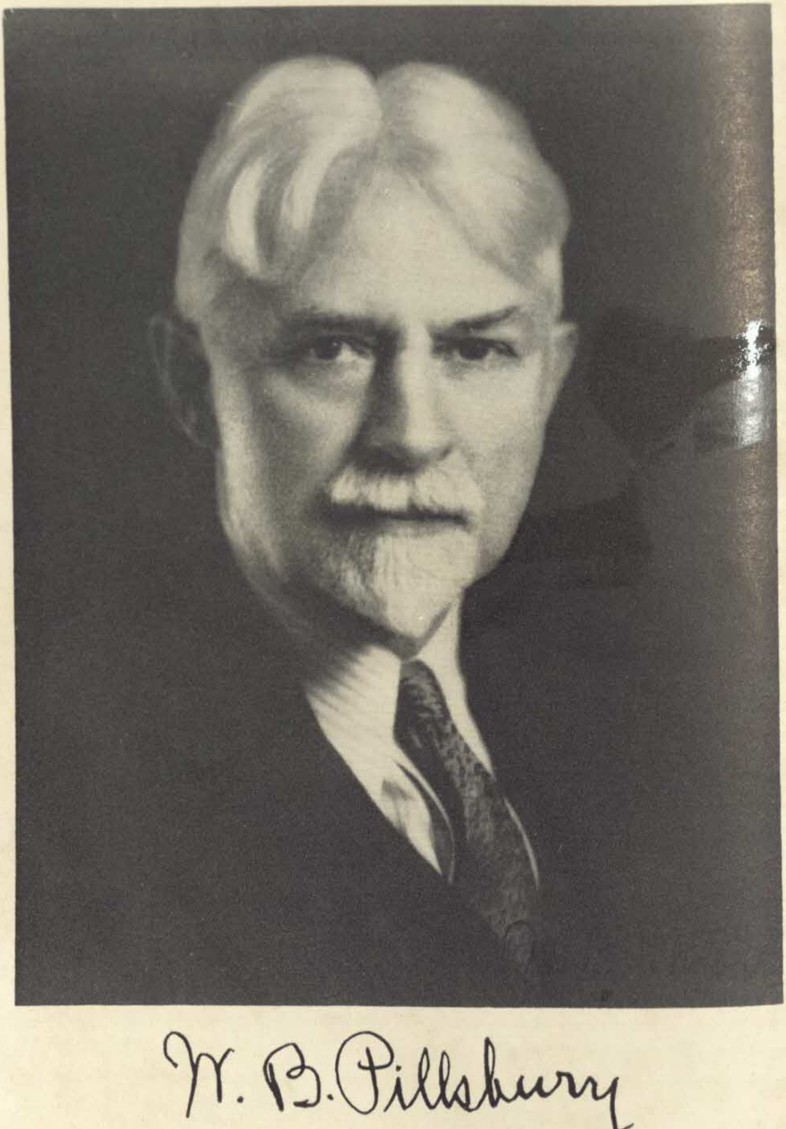
The chief contribution this book makes to psychology is to point out the importance of considering clothing whenever we are studying social phenomena and the psychology of clothing as an area for research.

MARY S. RYAN
Cornell University

Dreams and Personality Dynamics. Edited by M. F. DEMARTINO. Springfield, Ill., Charles C Thomas, 1959. Pp. xvii, 377. $10.50.

This is a useful collection of papers selected to acquaint a reader with some implications of dreams for the study of personality. Representative of its coverage is a discursive essay by Gardner Murphy on the function of the dream in the individual's psychological economy, a fascinating history of the paperback dream books which persist in the popular culture, actuarial surveys of the frequency of different forms of dreams as a function of demographhc characteristics, some studies of the personality correlates of various dream contents and several of the Dement papers on the psysiological concomitants of dreaming. With the exception of one article (by Medard Boss) translated from the German and a report of a doctoral dissertation, all the papers included have appeared previously. The appeal of the book is limited significantly by its appreciable cost and by the absence of a bibliography.

JACK BLOCK
University of California, Berkeley



W. B. Pillsbury

(See page 316)

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXXIV JUNE, 1961 No. 2

Walter Bowers Pillsbury: 1872–1960

Coöperating Editor 1897–1960

Walter Bowers Pillsbury, a coöperating editor of this JOURNAL for sixty-four years and one of the first generation of psychologists trained in America, died in Ann Arbor, Michigan, late in the morning of June 3, 1960. His death, due to a myocardial infarction, was sudden and unexpected as he had until then been enjoying excellent health and was at the time planning a trip to Europe to attend the meetings of the International Congress at Bonn, Germany. He was within a few days of his eighty-eighth birthday, as he was born on July 21, 1872, in Burlington, Iowa. He was the oldest of seven children—four boys and three girls—born to William Henry Harrison and Eliza Crabtree (Bowers) Pillsbury. His father was a Methodist minister who was sent by his Church, because of his administrative ability, through wide areas of Iowa and Nebraska.

The young Pillsbury was a precocious lad. He was graduated from high school before his sixteenth birthday and entered Penn College, Oskaloosa, Iowa, the following fall. His predilection toward psychology was aroused during his second year in high school by the chance reading of Carpenter's *Mental Physiology,* which he found in his father's library, to which he had free access. After reading this book, which he found to be most intriguing, he remarked to his father that he intended to specialize in psychology when he grew up.[1] Until late in his college course, however, this boyhood interest was not pursued as he found himself working toward law—not on his own initiative but rather from "acquiescence to family opinion."[2]

After his sophomore year at Penn College, Pillsbury transferred, in the fall of 1890, to the University of Nebraska. There his early enthusiasm for psychology was revived by a course that he elected under H. K.

[1] W. B. Pillsbury, Autiobiography, in *A History of Psychology in Autobiography,* Clark University Press, 2, 1932, 265. This is a strange autobiography, as he gives no information about his family, his early life and experiences—not even the date and place of his birth.

[2] *Idem,* 265.

Wolfe—Wundt's third American student and a magnetic teacher who turned a large number of his Nebraska students to careers in psychology. This course, which ran throughout the year, required considerable laboratory work and aroused in Pillsbury, as he afterwards reported, "a respect for experiment, a belief in a scientific psychology, and a desire to see thinking for its own sake more general."[3]

He was graduatel from Nebraska in 1892, and then taught "mathematics and a general assortment of school subjects" at Grand Island College, Grand Island, Nebraska. He remained here for but one year, as he accepted, in the spring of that year, the offer of a Sage Scholarship in Psychology at Cornell University, for which he had applied at Wolfe's suggestion. The fall of 1893 found him, therefore, in Ithaca, New York. His lifework was decided; his early ambition was to be achieved.

He found the academic atmosphere at Cornell congenial and stimulating. Titchener was beginning his second year there. Since he, like Wolfe, was a student of Wundt's there was little change in point of view. Experimentation, as with Wolfe, was the core of Titchener's teaching. In addition, however, Pillsbury found great pressure for publication which Titchener held was the end of a scholar's endeavor.

Psychology at Cornell was then, and for many years after, a part of the Sage School of Philosophy. As a Sage Scholar, it was Pillsbury's duty to acquaint himself with the trends in philosophy as well as in psychology and to write abstracts and book reviews for the *Philosophical Review,* a quarterly magazine that had shortly before (in 1891) been established by the Sage School. These duties assured the development of a facility in writing and, since many of the books received for review were French or German, in reading those foreign languages. In addition to these 'outside' duties and his courses in philosophy and psychology, Pillsbury served as an observer in the various studies being conducted in the laboratory. The exchange of observational hours among graduate students was standard procedure in the Cornell laboratory throughout Titchener's directorship of it. This practice provided the necessary observers, trained them in introspective reporting, and frequently gave them ideas for further research. Titchener, moreover, held that as much knowledge of psychology was gained by serving in a well-conducted experiment as in an academic course.

From his service as an observer in one experiment—Miss Washburn's doctoral dissertation on the influence of visual associations on the spatial perception of the skin[4]—Pillsbury obtained the idea for his first experi-

[3] *Idem,* 266.

[4] M. F. Washburn, Ueber den Einfluss der Geisichtsassociationen aud die Raumwahrnehmungen der Haut, *Philos. Stud.,* 11, 1895, 190-225.

mental research, which dealt with the problem of cutaneous localization with and without vision and with vision on life-sized photographs of the area (left forearm) stimulated.[5] He started on this 'minor' problem, as all researches performed by students except doctoral dissertations were designated, during the second semester of his first year—experimentation was indeed, as he soon found, the core of Titchener's instruction. Pillsbury's observers were Titchener, Miss Washburn, and three graduate students; so soon did he realize the benefits of the interchange of observational hours. He pushed the study rapidly forward, but withheld publication until Miss Washburn's study, which was temporally and logically prior, had appeared.

His scholarship at Cornell was renewed for a second year and he immediately turned to work upon his dissertation and, in collaboration with Titchener, upon the translation of Külpe's *Einleitung in die Philosophie.*[6] The topic for his dissertation was chosen in consultation with Titchener.[7] Pillsbury suggested several topics, which discussion revealed were unsuitable, and finally proposed an investigation of the mental processes involved in reading—an idea that came to him from the misreading of a street number on a letter box—which Titchener deemed appropriate. Pillsbury conceived of the problem as one in reading, but, as the investigation proceeded, it developed into a systematic study of apperception and attention, a turn that Titchener doubtless anticipated and directed.[8]

For his third year at Cornell, Pillsbury was advanced to an assistantship. During this year, he completed his dissertation, his translation of Külpe's *Einleitung,* a minor study on the projection of the retinal image,[9] and was awarded his Ph.D. degree. As teaching was, in those days, practically the only occupation open to a psychologist, and the demand for teachers in 1896 was small, Pillsbury found himself with a degree and no place to go. This situation, however, was soon rectified by his reappointment to the assistantship. The postdoctoral year proved highly profitable. He attended classes in physiology and related sciences in which he thought he was weak, rewrote his dissertation for publication, read the galley proofs of his translation, and devoted himself to teaching as he,

[5] Pillsbury, Some questions of the cutaneous sensibility, this JOURNAL, 7, 1895, 190-225.

[6] Oswald Külpe, *Introduction to Philosophy,* Translated by W. B. Pillsbury and E. B. Titchener, 1897, x. 256.

[7] In Pillsbury's day, Titchener permitted his students to propose the topics of their dissertations and he selected one from among them. In later days, he proposed and the students selected; a preferred procedure as he knew, better than they, what was pertinent and timely.

[8] Pillsbury, A study in apperception, this JOURNAL, 8, 1897, 315-393.

[9] Pillsbury, The projection of the retinal image, this JOURNAL, 9, 1897, 56-60.

though still an assistant, was permitted to give independently a course in the Department.

His dissertation and his translation, both published in the spring of 1897, attracted attention to him, and calls to positions in psychology were now not lacking. Among those received was an instructorship at the University of Michigan. Though Cornell met this offer,[10] Pillsbury accepted Michigan's invitation because it entailed the welcome responsibility of establishing a new laboratory. At the same time, he was invited by G. Stanley Hall to join the Coöperating Board of Editors of this JOURNAL—the first of Titchener's students so honored. Both of these affiliations were maintained by Pillsbury throughout the duration of his long life.

Psychology at Michigan was then taught in the Department of Philosophy, of which R. M. Wenley was the head. Pillsbury, however, was given full charge of the laboratory, which he quickly established, and all of the work in psychology. His advancement, slow according to present standards, was then considered to be rapid as he was promoted to an assistant professorship in 1900, to a junior professorship and the directorship of the psychological laboratory in 1905, and to a professorship in 1910. Psychology was continued in the Department of Philosophy, however, until Wenley's death in 1929. Then a separate Department of Psychology was created and Pillsbury was made chairman, the position he held until June of 1942, when he was retired at the age of seventy years to an emeritus professorship.[11]

Pillsbury carried to Michigan his enthusiasm for research and publication gained from Titchener. His early years there were marked by hard work. During one semester, he had, as he reported in his autobiography, "a teaching schedule of forty-two hours a week."[12] For several years after joining the staff at Michigan he further improved his knowledge of the biological sciences, begun during the postdoctoral year at Cornell, by taking courses in neural anatomy and by collaborating with W. P. Lombard, professor of physiology, on studying the influence of circulatory and respiratory changes, in particular the Traube-Hering waves, on the fluctua-

[10] An announcement of Pillsbury's appointment to an instructorship in experimental psychology at Cornell appears in the April, 1897 number of this JOURNAL, (9, 1897, 430).

[11] Pillsbury was criticized, particularly during the early and middle years of his service at Michigan, for not obtaining more rapid promotions of his colleagues in psychology. He, however, was not at fault, as Wenley, a strong and positive character, was the administrative head of the Department. If a promotion were available, a philosopher, to whose work Wenley stood close, was apt to receive it. Pillsbury was a mild, gentle man; not a fighter, as the long union of psychology and philosophy at Michigan attests.

[12] *Op. cit.*, 272.

tion of attention. Two papers published jointly with Lombard and nine papers from his laboratory resulted from this excursion into physiology.[13]

In 1902, the beginning of the most active decade of his life, he revised his translation of Külpe's *Einleitung*. The manuscript of the revision was, however, lost when the English firm to which it had been sent for publication went into bankruptcy shortly after receiving it. Due to the time required for the firm's reorganization and in the preparation of a new manuscript, publication was delayed until 1904. In the meantime, in 1903, shortly before he left for a trip abroad to visit the European laboratories in psychology, he wrote his first book. This book, which was on attention, grew out of his dissertation and the experimental work on attention done in his and Lombard's laboratories. It was written in English but translated into French and published by a Paris firm as no American or English firm was willing to undertake the risk involved in the publication of such a technical and highly specialized work. Because of the time required for the translation, the publication of *L'Attention* was delayed until 1906. Except for the translation, which was evidently not made by a French psychologist as it was not always clear nor correct, the book was well received.[14] Soon offers to publish the English text were forthcoming. Pillsbury chose, however, to rewrite and greatly to enlarge it, with the consequence that the English edition of *Attention,* which appeared in 1908, was practically a new book. Of all his numerous works, it is probably the one for which he is best known.

During the first semester of the academic year 1908–1909, he gave a series of lectures on "The Psychology of Reasoning" at Columbia University, a topic of interest acquired at Cornell while taking a course on "the modern logicians" with J. E. Creighton, and sustained over the intervening years by a series of articles on reason, thought, and judgment.[15] These leceures were expanded and published under their title in 1910[16]—the year that he held the presidency of the American Psychological Association, was promoted to his professorship at Michigan, joined *The Psychological Review* as Advisory Editor, and also the year marked by the trans-

[13] W. P. Lombard and W. B. Pillsbury, A new form of piston recorder and some of the changes it records, *Amer. J. Physiol.,* 3, 1899, 186-200; Secondary rhythm of the normal heart beat, *idem,* 201-226. For the papers from his laboratory, see "Studies from the Psychological Laboratory of the University of Michigan" under III. Editorial, in his Bibliography, *infra,* p. 174.

[14] Cf. review by M. F. Washburn, this JOURNAL, 17, 1906, 593-596, esp. 596.

[15] Pillsbury, Psychological nature of causality, *Philos. Rev.,* 13, 1904, 409-419; An attempt to harmonize current psychological theories of judgment, *Psychol. Bull.,* 15, 1908, 150-157.

[16] *The Psychology of Reasoning,* 1910, ix + 306.

lation of the English edition of his *Attention* into Spanish and its publication in Madrid.

From the vantage point of these achievements, he turned from the laboratory to the writing of textbooks and non-experimental articles. His first elementary textbook, *The Essentials of Psychology,* published in 1911, was eclectic in its point of view. Pillsbury was not a systematist, either by inclination or training, and he never pledged allegiance to any school.[17] He moved freely, therefore, from one system to another as his interest in the topics being considered dictated. Though professedly *The Essentials* was written from a functional point of view, he made use "of the results of structural psychology wherever they throw light upon function or are interesting for themselves."[18] The catholicity of his treatment proved to be popular. The book was widely adopted. It went through numerous printings, a revised second edition in 1920, and it was still so greatly in demand that it was profitable to publish a third edition in 1930.

The Essentials was written for a short, a single semester, course. It was not adequate, however, for the longer, year-courses in elementary psychology which, by then, universities were offering in increasing numbers. Michigan very early offered two elementary courses: a semester course for students in Education and for those who wished merely an acquaintanceship with psychology; and a long, two-semester course for students intending to major in psychology and for those desiring a more thorough survey than could be obtained in the short course. To meet the needs of the long course, Pillsbury wrote *The Fundamentals of Psychology,* which was published in 1916. This book is also eclectic in its point of view. As stated in the Preface, Pillsbury drew in its writing upon "the work of all schools": functionalism, structuralism, and behaviorism—the latter had just recently appeared upon the psychological scene in America. He subordinated theory to fact and expounded the latter from the point of view from which it could most favorably be treated. Since beginning students have little or no appreciation of what is involved in a shift in point of view, the lack of system was advantageous and soon this book had achieved the popularity of his first. It, too, was widely adopted and used for many years: a revision was published in 1922, and a third edition in 1934.

---

[17] Pillsbury left Cornell before the distinction between structural and functional psychology was drawn. The center of interest while he was there was in the laboratory. Titchener did not identify himself as a structural psychologist until several years later when he published the following articles: The postulates of a structural psychology, *Philos. Rev.,* 7, 1898, 449-465; Structural and functional psychology, *ibid.,* 8, 1899, 290-299.

[18] *Op. cit.,* vii.

During World War I, work was heavy upon Pillsbury. He remained at his post and took over the load of his colleagues as they entered the Services and other work for the National defense. There was little time for independent work. He bethought himself, however, of the observations that he had made in Greece during his second trip abroad in 1912–1913, on mixed allegiance. To assist in the Balkan-Turkish War being waged at that time, many Greco-Americans returned to their mother country. Why did they leave America to fight for the country from which they had migrated? The answer seemed to Pillsbury to lie in the shift from nationality to internationalism. He developed this theme in *The Psychology of Nationality and Internationalism,* published in 1919.

Beginning with his *L'Attention* in 1906, Pillsbury published a new book in psychology at about the rate of one every three years. In addition to the six mentioned above, he published the revised edition of his *Essentials* in 1920; the revised edition of his *Fundamentals* in 1922; *Education as the Psychologist Sees It* in 1925; *The Psychology of Learning* (with C. L. Meader) in 1928; *The History of Psychology* in 1929; the third revised edition of his *Essentials* in 1930; *An Elementary Psychology of the Abnormal* in 1932; the third edition of his *Fundamentals* in 1934; the second edition of his *History* in 1937; and his last work, *A Handbook of General Psychology* (with L. A. Pennington) in 1942, the year of his retirement.

The break in the three-year rhythm of new books, which occurred in 1922, was doubtless due to his trip to France, where he spent the academic year 1921–1922 as an exchange professor. He lectured one term at Toulouse, the second at Montpellier, and the third at the Sorbonne, in Paris, and, as opportunity permitted, toured the country, lecturing at the smaller universities. In this schedule, little time availed for the writing of a book. He did, nevertheless, as he states in his autobiography,[19] find time to begin the writing of his provocative book on *Education as the Psychologist Sees It.* In this book he defended the theses that the results of education are actually the results of selection; that only the students with the greatest native ability succeed in passing through the educational system; and that they would have been successful men or women whether they passed through it or not.

The Psychology of Language, which followed in 1928, was the outgrowth of a joint course he had given for ten years with C. L. Meader, professor of linguistics. His *History of Psychology,* undertaken to fill a real need, as no American had written a text in this field, was also an outgrowth

[19] *Op. cit.,* 289.

of a course that he had taught for many years. He wrote it in Berlin during the summer of 1928, where he had gone to gain first hand knowledge of Gestaldt psychology and to be with his daughter who was studying there at that time. Though his book was one of three on History that appeared in 1929; his was sufficiently successful, because it was brief and simply written, to warrant the publication of a second edition in 1937.

His last book, *A Handbook of General Psychology,* written with L. A. Pennington, one of his former students, is a strange potpourri of psychological information, as its subtitle, "A Summary of Essentials and A Dictionary of Terms," indicates. It is divided into four parts: Part I is a brief elementary textbook; Part II, a 73-page dictionary of psychological terms; Part III contains 105 brief biographical sketches of psychologists; and Part IV is a bibliography of 62 reference books in general psychology. It is difficult to see for 'what part of the trade' it was written.

In addition to these books, Pillsbury wrote many articles on psychology and cognate subjects. During his early years, these articles were concerned chiefly with experiments; his middle years, with essays and timely topical reviews; and his latter years, with historical and necrological notes. Due to his long life, he was frequently called upon to mark the passing of his early contemporaries (Titchener, Washburn, Claparède, Jastrow, Yoakum, Cattell, Carr).

Eleven books—two of them going through three editions, one through two editions, and one appearing in two foreign languages—one translation which went through two editions, and sixty-nine articles secure him place in the history of American psychology. He was one of the pioneers who founded laboratories and departments, who wrote books and edited journals, who prepared the way for the generations of psychologists that followed him. He received, in recognition of his achievements, all the honors within the power of his confrères to bestow upon him. He was elected to the presidency of the American Psychological Association in 1910; to the presidency of the Western Philosophical Association in 1907; to the vice-presidency and chairmanship of the psychological section of the American Association for the Advancement of Science in 1913; to membership in the National Academy of Science, in 1925; to the National Research Council in 1921; as a foreign associate of the *Société française de psychologie* in 1925; and he held membership in the Linguistic Society of America, The Society of Experimental Psychologists, and the Midwestern Psychological Association. In addition to the lectureships at Columbia University and in France, previously mentioned, he gave the Henry Russel lecture at the

University of Michigan in 1933. In that same year his Alma Mater, the University of Nebraska, conferred the LL.D degree upon him.

Throughout his professional career, Pillsbury gave freely of his time and experience to editorial work. In addition to his long service (sixty-four years) on the editorial board of this JOURNAL—the longest editorial service on record—he served *The Psychological Review*, as noted above, for twenty years, from 1910–1929. He was also an associate editor of *The Journal of Social Psychology* for thirty-one years, joining its staff at its founding in 1930 and remaining on it until his death. Earlier, he edited the series of "Studies from the Psychological Laboratory of the University of Michigan," which ran from 1901 to 1905; and, with J. W. Baird and M. F. Washburn, the commemorative volume, *Studies in Psychology*, published in 1917 in Titchener's honor by his colleagues and former students.

Pillsbury was subject to *petit mal* through most of his life.[20] This affliction did not, as we have seen, affect his productivity nor his ability to work, nor did it shorten the span of his life, but it may have affected his personality. He was not the self-assertive man that one of his position, honors, and achievements might well have been. He was a gentle man—modest, retiring, and unobtrusive. Though always cordial and friendly with acquaintances, he held himself a bit aloof and allowed himself but few close and intimate friendships. Those who succeeded in breaking through his reserve, found his friendship firm and true.

He is survived by his wife, Mrs. Margaret M. (Millbank) Pillsbury; daughter, Margaret Elizabeth (Mrs. Warren Baxter) and her three children and four grandchildren; and son, Walter Millbank Pillsbury.

KARL M. DALLENBACH
University of Texas

[20] Despite this infirmity, Pillsbury was mentally alert and in excellent physical condition throughout his life. His senses, especially his eyes, showed no deterioration with advancing age. When he passed middle life, he needed glasses for close work, but not for other activities. His eyes were exceptional, without detectable astigmatism. His corneas were in such excellent condition that he was asked to will his eyes to the eye-bank of the University Hospital. He agreed, and within a few hours after his death his corneas were transplanted, restoring the vision of a fellow man. Thus his outstanding characteristics, kindness and humanity, were expressed even after death. His body, at his request, was cremated.

A BIBLIOGRAPHY OF WALTER BOWERS PILLSBURY

By M. C. McGrade and Karl M. Dallenbach, University of Texas

I. Books

(1) *L'Attention*. Paris, Doin, 1906, 305.
(2) *Attention*. London, Swan, Sonnenschein; New York, Macmillan, 1908, xi + 345.
(3) *Same*. Spanish edition. Madrid, June, 1910, viii + 343.
(4) *The Psychology of Reasoning*, New York, Appleton, 1910, 306.
(5) *The Essentials of Psychology*. New York, Macmillan, 1911, xi + 358.
(6) *Same*. Revised edition. New York, Macmillan, 1920, ix + 466.
(7) *Same*. Third edition. New York, Macmillan, 1930, ix + 466.
(8) *The Fundamentals of Psychology*. New York, Macmillan, 1916, ix + 562.
(9) *Same*. Revised edition. New York, Macmillan, 1922, xvi + 590.
(10) *Same*. Third edition. New York, Macmillan, 1934, xii +663.
(11) *The Psychology of Nationality and Internationalism*. New York, Appleton, 1919, viii + 314.
(12) *Education as the Psychologist Sees It*. New York. Macmillan, 1925, vii + 342.
(13) (With C. L. Meader). *The Psychology of Language*. New York, Appleton, 1928, viii + 306.
(14) *The History of Psychology*. New York, Norton; London, Allen and Unwin, 1929, 326.
(15) *Same*. Second edition. Ann Arbor, George Wahr, 1937, 326.
(16) *An Elementary Psychology of the Abnormal*. New York, McGraw-Hill, 1932, 375.
(17) (With L. A. Pennington). *Handbook of General Psychology: A Summary of Essentials and a Dictionary of Terms*, New York, Dryden Press, 1943, xiv + 400.

II. Translations

(1) (With E. B. Titchener). *Introduction to Philosophy*, by Oswald Külpe, [Translation of *Einleitung in die Philosophie*. London, Swan, Sonnenschein, 1895.] New York, Macmillan, 1897, x + 256.
(2) *Same*. Revised edition. London, Sonnenschein, 1904, x + 256.

III. Editorial

(1) Coöperating editor, *The American Journal of Psychology*, from 1897 (Vol. 9) to 1960 (Vol. 73).
(2) Editor, Studies from the Psychological Laboratory of the University of Michigan.
   I. J. W. Slaughter, The fluctuations of the attention in some of their psychological reactions, *Amer. J. Psychol.*, 12, 1901, 313-334;
   II. R. W. Taylor, The effect of certain stimuli upon the attention wave, *Amer. J. Psychol.*, 12, 1901, 335-345.
   III. W. B. Pillsbury, Does the sensation of movement originate in the joint? *Amer. J. Psychol.*, 12, 1901, 346-353.
   IV. J. W. Bair, Development of voluntary control, *Psychol. Rev.*, 8, 1901, 474-510.
   V. H. C. Stevens, The relation of the fluctuations of judgments in the estimation of time intervals to vaso-motor waves, *Amer. J. Psychol.*, 13, 1902, 1-27.
   VI. W. B. Wright, The relation between the vaso-motor waves and reaction time, *Psychol. Rev.*, 11, 1904, 179-185.

VII. C. E. Galloway, The effect of stimuli upon the length of Traube-Hering waves, *Amer. J. Psychol.*, 15, 1904, 499-512.
VIII. Bertha Killen, The effects of closing the eyes upon the fluctuation of the attention, *Amer. J. Psychol.*, 15, 1904, 512-514.
IX. Bernice Barnes, Eye movement, *Amer. J. Psychol.*, 16, 1905, 199-207.

(3) Advisory editor, *The Psychological Review*, from 1910 (Vol. 17) to 1929 36).
(4) Co-editor, *Studies in Psychology: Titchener Commemorative* Volume, 1917, ii + 337.
(5) Associate editor, *The Journal of Social Psychology*, from 1930 (Vol. 1) to 1960 (Vol. 51).

IV. Articles, Notes, Discussions

(1) Some questions on the cutaneous sensibility, *Amer. J. Psychol.*, 7, 1895, 42-57.
(2) The projection of the retinal image, *Amer. J. Psychol.*, 8, 1897, 250-267.
(3) A study in apperception, *Amer. J. Psychol.*, 8, 1897, 405-413.
(4) (With W. P. Lombard), A new form of piston recorder and some of the changes it records, *Amer. J. Physiol.*, 3, 1899, 186-200.
(5) (With W. P. Lombard), Secondary rhythms of the normal human heart, *Amer. J. Physiol.*, 3, 1899, 201-228.
(6) Does the sensation of movement originate in the joint?, *Amer. J. Psychol.*, 12, 1901, 346-353.
(7) Attention waves as a means of measuring fatigue, *Amer. J. Psychol.*, 14, 1903, 541-552.
(8) The psychological nature of causality, *Philos. Rev.*, 13, 1904, 409-419.
(9) A suggestion toward a re-interpretation of introspection, *J. Philos., Psychol. & Sci. Methods*, 1, 1904, 225-228.
(10) An apparent contradiction in the modern theory of judgment, *J. Philos., Psychol. & Sci. Methods*, 2, 1905, 568-573.
(11) Trial and error as a factor in evolution, *Pop. Sci. Mo.*, 68, 1906, 276-282.
(12) The ego and empirical psychology, *Philos. Rev.*, 16, 1907, 387-407.
(13) An attempt to harmonize the current psychological theories of the judgment, *Psychol. Bull.*, 4, 1907, 237-242.
(14) The effects of training on memory, *Educ. Rev.*, 36, 1908, 15-27.
(15) Meaning and image, *Psychol. Rev.*, 15, 1908, 150-157.
(16) The role of type in simple mental processes, *Philos. Rev.*, 20, 1911, 83-99.
(17) Attention and interest, *Psychol. Bull.*, 8, 1911, 169-172.
(18) The Innsbruck Congress of 1910, *Psychol. Bull*, 8, 1911, 207-212.
(19) The place of movement on consciousness, *Psychol. Rev.*, 18, 1911, 83-99. (Presidential address, A. P. A., Minneapolis Meeting, December, 1910.)
(20) Rousseau's contribution to psychology, philosophy, and education, *Pop. Sci. Mo.*, 81, 1912, 331-335.
(21) Attention and interest, *Psychol. Bull.*, 8, 1912, 193-196.
(22) "Fluctuation of attention" and the refractory period, *J. Philos., Psychol. & Sci. Methods*, 10, 1913, 181-185.
(23) Attention, *Psychol. Bull.*, 11, 1914, 166-169.
(24) Attention, *Psychol. Bull.*, 12, 1915, 177-178.
(25) The function and test of definition and method in psychology, *Sci.*, 41, 1915, 371-389.
(26) The mental antecedents of speech, *J. Philos., Psychol. & Sci. Method*, 12, 1915, 116-127.
(27) Attention and interest, *Psychol. Bull.*, 13, 1916, 205-206.
(28) The new development in psychology in the past quarter century, *Philos. Rev.*, 26, 1917, 56-69.
(29) Attention and interest, *Psychol. Bull.*, 14, 1917, 169-170.
(30) The datum, in *Philosophical Essays in Honor of James Edwin Creighton*, New York, Macmillan, 1917, 162-174.
(31) Principles of explanation in psychology, in *Studies in Psychology: Titchener Commemorative Volume*, Worcester, Wilson, 1917, 18-29.

(32) Attention and interest, *Psychol. Bull.*, 15, 1918, 108-111.
(33) Attention, *Psychol. Bull.*, 16, 1919, 248-249.
(34) Attention and interest, *Psychol. Bull.*, 17, 1920, 259-260.
(35) Selection: An unnoticed function of eduction, *Sci. Mo.*, 12, 1921, 62-74.
(36) Suggestions for a compromise of existing controversies in psychology, *Psychol. Rev.*, 29, 1922, 250-266.
(37) Le dévelopment de la psychologie en Amérique, *J. de Psychol.*, 7, 1923, 597-613.
(38) L'épreuve et l'erreur, loi des opérations mentales, *Rev. de Phil.*, 96, 1923, 203-213.
(39) (With C. H. Griffith), An experiment on indirect measures of fatigue, *Proc. 7th Int. Cong. Psychol.*, 1924, 212-215.
(40) Recent naturalistic theories of reasoning, *Scientia*, 36, 1924, 23-32.
(41) The present status of psychology in the United States, *Scand. Sci. Rev.*, 3, 1925, 199-208.
(42) *Gesalt* vs. concept as a principle of explanation in psychology. *J. abn. soc. Psychol.*, 21, 1926, 14-18.
(43) What is native in the so-called instincts?, *Amer. J. Psychol.*, 39, 1927, 42-53.
(44) The psychology of Edward Bradford Titchener, *Philos. Rev.*, 37, 1928, 95-108.
(45) The utility of emotions, in M. L. Reymert (ed). *Feelings and Emotions: The Wittenberg Symposium*, 1928, 116-123.
(46) The cinematoscope as a memory apparatus, *Amer. J. Psychol.*, 41, 1929, 649-650.
(47) Autobiography, in Carl Murchison, (ed.), *A History of Psychology in Autobiography*, 2, 1932, 265-296.
(48) The Wundt centenary, *Amer. J. Psychol.*, 45, 1933, 176.
(49) The unit of experience—meaning or Gestalt, *Psychol. Rev.*, 40, 1933, 481-497.
(50) A case of hemeralopia, *Amer. J. Psychol.*, 46, 1934, 655-658.
(51) A case of mental conflict, *J. abn. soc. Psychol.*, 31, 1936, 312-323.
(52) Body form and success in studies, *J. soc. Psychol.*, 7, 1936, 129-139.
(53) (With B. R. Schaefer), A note on 'advancing and retreating' colors, *Amer. J. Psychol.*, 49, 1937, 126-130.
(54) Correlation between the university grades of parents and children and of siblings, *Amer. J. Psychol.*, 52, 1939, 634-638.
(55) Retroactive inhibition tested immediately, *Psychol. Bull.*, 36, 1939, 544 (abstract).
(56) (With Anne Sylvester), Retroactive and proactive inhibition in immediate memory, *J. exp. Psychol.*, 27, 1940, 532-545.
(57) Margaret Floy Washburn: (1871-1939), *Psychol. Rev.*, 47, 1940, 99-109.
(58) Edouard Claparède, 1873-1940, *Psychol. Rev.*, 48, 1941, 271-278.
(59) (With H. L. Raush), An extension of the Köhler-Restorff inhibition phenomenon, *Amer. J. Psychol.*, 56, 1943, 293-298.
(60) Titchener and James, *Psychol. Rev.*, 50, 1943, 71-73.
(61) Propaganda and the democratic state, *Sci. Mo.*, N.Y., 56, 1943, 549-555.
(62) Joseph Jastrow: 1863-1944, *Psychol. Rev.*, 51, 1944, 261-268.
(63) Clarence Stone Yoakum: 1879-1945, *Psychol. Rev.*, 53, 1946, 195-198.
(64) James McKeen Cattell, 1860-1944, *Nat. Acad. Sci. Biogr. Mem.*, 25, 1947, 1-16.
(65) Attention, in P. L. Harriman, (ed.), *Encyclopedia of Psychology*, 1947, 46-48.
(66) Knowledge in modern psychology, *Psychol. Rev.*, 57, 1950, 328-333.
(67) The concept in psychology, *Année Psychol.*, 50, 1951, 97-104.
(68) Harvey A. Carr: 1873-1954, *Amer. J. Psychol.*, 68, 1955, 149-151.
(69) Errata, *Amer. J. Psychol.*, 68, 1955, 325.

FORM AND ITS ORIENTATION: A CHILD'S-EYE VIEW

By Lila Ghent, The George Washington University

The orientation of pictorial material usually is determined by the position of the figure with respect to a frame of reference provided by the environment or by the observer. The figure is considered right side up, or correctly oriented, when it is in the usual or familiar position with reference to the frame. Nonrealistic, or geometric, figures are not, however, usually considered to have a 'right side up' orientation (except in certain esthetic judgments).

Some chance observations indicated that young children show preferences for the orientation of geometric forms, *i.e.* they consider certain nonrepresentational forms to be right side up in one orientation and upside down in another. Such a finding would be surprising in any age-group, but it is particularly unexpected in preschool children in view of the customary belief that young children are unresponsive to the orientation of forms. This report describes a systematic investigation of these curious preferences for orientation and discusses the implicatons of the results for the more general problem of the influence of orientation on the child's perception of form.

Experiment I

Children of various ages were tested to determine the consistency of preference at different ages, and various types of form were used in an attempt to analyze the aspects of the stimulus that determined preference.

Subjects. The *Ss* were 78 children between the ages of 4-8 yr. There were 22 aged 4 yr., 26 aged 5 yr., 14 aged 6 yr., and 16 aged 7-8 yr., with approximately the same number of boys and girls in each group. The children were enrolled in a child care center situated in a low-income area in New York.[1]

Procedure. Twenty-six pairs of pictures were presented to each child, and he was asked to point to the one that was upside down or wrong. The members of each

---

* Received for publication December 4, 1959. This study was begun at the Psychophysiological Laboratory, New York University-Bellevue Medical Center, and was supported by a grant to the laboratory from the Commonwealth Fund of New York. The work was continued at the Pediatrics Department of the State University of New York, Downstate Medical Center, and was supported in part by grant M-2444 from the National Institute of Mental Health, Public Health Service.

[1] The writer is indebted to the Director and teachers of the Bronx River Child Care Center for their coöperation.

pair of pictures were identical except that one was rotated 180° with respect to the other. Five pairs of realistic forms were presented first (rooster, flowerpot, shoe, cat, chair), then the 16 pairs of the nonrealistic forms shown in Fig. 1, and finally another 5 pairs of realistic forms (tree, boat, cup, clown, horse). The series was always shown in the same order.

E sat beside *S* and held two pictures of every form with vertical axes reversed (one in each hand) in front of *S*, with a distance of 4–6 in. between the closer edges of the cards. *S* was allowed as much time as he wished to make a choice and was allowed to change his choice, if he desired. He was not told, however, whether his choices were right or wrong, although he was encouraged in a rather general way throughout the testing procedure. The only exceptions were the very few instances in which *S* made an error on the realistic forms. When this happened, *E* said, "Are you sure that is the upside down one? Go slowly now, and always pick the one that looks upside down or wrong." The *S*s of 6 yr. and older frequently balked at choosing between the nonrealistic forms, but *E* urged them to guess.

Results. Perhaps the most striking result was the ready acceptance by the preschool children of the problem of choosing the upside down member of a pair of nonrealistic forms. When the pairs of nonrealistic forms were introduced, the *S*s of 4 and 5 yr. made their choices with the same speed and ease that they had shown with the representational forms. They did not hesitate, nor did they comment on the pictures at this point; in general, they behaved as though the problem had not changed. The children over 6 yr. of age usually hesitated for a few seconds when the nonrealistic forms were first presented, and they often were reluctant to choose between the cards. These older children would say, "I can't tell," or "It doesn't matter," or the like.

A summary of the responses to each pair of nonrealistic forms is presented in Fig. 1, which shows the percentage of children in each age-sex grouping choosing the card as upside down in the orientation shown here. In the 4-yr.-old group, the frequency of choosing the card as upside down in the orientation shown in Fig. 1 was significantly greater than would be expected by chance (5% level) for all cards except Forms 4, 7, 13, 14, and 16. The consistency of the response to some items was extraordinary in the youngest group, with more than 95% of the 4-yr.-olds choosing as upside down Forms 1, 3, 5, 9, 10, and 15.

With increasing age, the choices remained essentially the same, and significantly different from chance in all age-groups (with a minor exception) for Forms 5, 9, 10, 11, and 12. (The exception is that the responses of the *S*s of 7–8 yrs. fell just short of the 5% level of significance for Forms 11 and 12.) Responses that were consistent in the 4-yr.-old group, but random in the oldest group, occurred for Forms 6, 2, 3, and 15, with significant differences ($p < 0.02$) appearing between the youngest and oldest groups for the latter three items. Other items showed a change from ran-

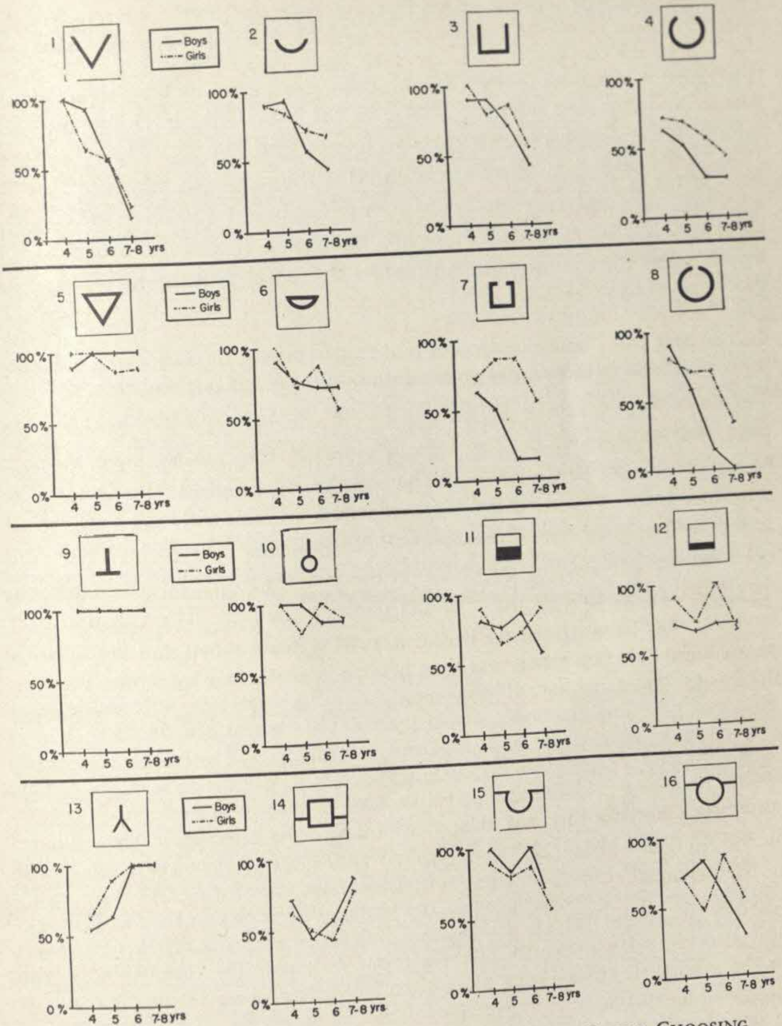


FIG. 1. PERCENTAGE OF BOYS AND GIRLS IN EACH AGE-GROUP CHOOSING CARD AS UPSIDE DOWN IN THE ORIENTATION SHOWN.
Each card was 4 in. on a side and did not have the black border shown here.

domness to consistency (Forms 7, 13) or an actual reversal of preference for orientation (Forms 1, 8).

When the incomplete circle (Form 8) was oriented with the gap on top, the card was considered upside down by 86% of the 4-yr.-old *S*s, whereas it was considered right side up by 81% of the 7–8-yr.-old *S*s ($p < 0.001$).

The responses to the incomplete square (Form 7) were random in the 4-yr.-old group, but the older boys consistently preferred the square with the gap on top. Comparison of the responses to the incomplete square of the 4- and 5-yr.-old boys with those of the 6–8-yr.-old boys showed a significant difference between the age-groups ($p < 0.02$).

Both for the incomplete circle and for the square, a sex-difference appeared in the older age-groups. Comparison of the responses of the boys and girls in the 6–8-yr.-old groups showed a significant difference between them for the incomplete square ($p < 0.01$) and the incomplete circle ($p < 0.05$).

It is clear that consistent responses to the orientation of form appear, not only for realistic figures, but also for geometric forms, in the 4- and 5-yr.-old child. Do these results reflect a true preference for orientation of nonrealistic forms, or is the preference due to the child's perception of the geometric forms as representations of real objects? The second interpretation is not supported by the children's spontaneous verbalizations, nor by their responses to questions. For example, the children did not consistently name, or say that they recognized, the card considered right side up. Sometimes, in fact, a child would give two names for each form, one for each orientation. (The crescent was called a bridge in one orientation and a bowl in the other by the same *S*, who nevertheless chose the 'bowl' as upside down.) When the young *S*s were asked how they knew which picture was wrong, they usually replied, "I see it" or "I can tell" or the like. It is striking that not a single child suggested that the right side up picture was right because he could see what it was.

Granted that a genuine, although unexplained, preference for orientation exists, let us consider this phenomenon from the point of view of age. The frequency of preferences was highest in the *S*s of 4 and 5 yr. of age, and essentially the same in both age-groups. There was a decline in preferences for orientation with age, but this change cannot be described as a simple decrease in strength of preference. A preference, when it appeared in the older *S*, tended to be as strong as in the younger *S* (Forms 5, 9, 10, 11, 12); fewer figures, however, evoked preferences in the older *S*. Perhaps the most puzzling change with age was the appearance of preferences that were not present in the young child. The changes in response to the 'V' (Form 1) and the 'Y' (Form 13) are not surprising, since the responses of the older *S*s are consistent with the perception of these forms as letters. The changes in response to the incomplete circle (Form 8) and the incomplete square (Form 7) cannot be accounted for so simply.

An interesting aspect of the change with age in response to the figures with a gap is that the change occurred earlier in the boys than in the girls. An isolated finding of different responses to visual form by boys and girls would be of little significance in itself. It also has been found, however, that, in the same age-group, boys are better than girls in the recognition of tachistoscopically presented realistic forms.[2] The finding of two instances in which the response to form can be considered to be more developed in boys than girls is of significance for understanding sex-differences

---

[2] Lila Ghent, Recognition by children of realistic figures presented in various orientations, *Canad. J. Psychol.*, 14, 1960, 249-256.

in perception and raises the possibility that these changes (with age) in perception reflect related processes.

The most important problem raised by the finding of preferences for orientation is that of defining the characteristics of the cards selected as upside down. For this discussion, changes with age will be ignored, and the responses of the preschool Ss will form the basis for the analysis. Some of the forms were originally designed on the hypothesis that 'openness' in the top part of the figure characterized the upside down card. While most of the forms with an open top were considered upside down, the 'openness' did not appear to be essential, since the same forms with a partially or completely closed top also were considered wrongly oriented (compare, for example, Forms 1–4 and 5–8 of Fig. 1). Another possibility is that the upside down cards were unstable or top heavy, but the data do not support this view either (see Forms 9–12, 14, and 16).

The interpretation that appears to be most consistent with the data is that the card was upside down when the focal portion was in the lower half of the card. That is, let us assume that the figures eliciting significant preferences had one portion that caught the 'attention,' or drew the eye, more readily than did other portions. The angle of the *V*, the point of intersection of the *T*, the rounded portion of the crescent, and so on, each might be considered the focal portion of the particular form. When these portions of the figures were at the bottom, the cards were called upside down by the children.

Experiment II

The main difficulty with this interpretation is that the designation of 'focal' is made *post hoc*. For the experiment to be described now, a set of very simple cards was designed in which the focal part was designated beforehand, and, in most instances, there was little ambiguity as to whether or not the stated portion could be described as focal (Fig. 2). These simple cards were used to test the hypothesis that the cards were considered upside down when the focal portion was in the lower part. On the assumption that the hypothesis would be confirmed, an additional set of cards was included in the second experiment to explore some determinants of the focal portion of complex figures—brightness, size, and other variables.

Subjects. Forty children of preschool age were tested in the nursery school previously described, but none of these children had participated in the first experiment. The 4-yr.-old group consisted of 11 boys and 11 girls, and the 5-yr.-old group of 9 boys and 9 girls.

Procedure. Thirty-four pairs of cards were used, 10 realistic, and 24 nonrealistic forms. One member of each pair of the nonrealistic forms can be seen in Fig. 2 and Fig. 3. The realistic forms were the same as those previously used, but, for this experiment, one pair appeared at the beginning and at the end, and the remaining eight pairs were distributed throughout the series. The cards were presented in one

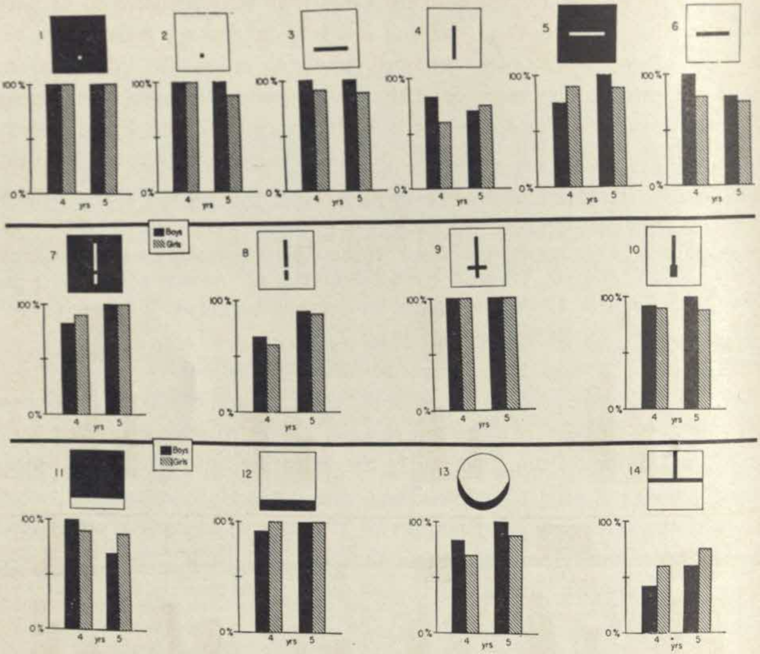


FIG. 2. PERCENTAGE OF BOYS AND GIRLS IN EACH AGE-GROUP CHOOSING CARD AS UPSIDE DOWN IN THE ORIENTATION SHOWN.
With the exception of Card 13, each card was 4 in. on a side and did not have the black border shown here.

sequence to half the children in each age-sex grouping, and in the reverse order to the remaining children in each group.

Results. Fig. 2 shows the responses to the simple cards in which there was little ambiguity, in most instances, as to which portion constituted the focal part. The preferred position of a simple line or dot was investigated with Cards 1–6. It is clear that the card was considered upside down when the figure was in the lower part of the card. When the line was centered, it was considered wrong when placed horizontally instead of vertically. The next set of cards shows that, with a simple line interrupted in various ways, the cards were considered upside down when the point of interruption (the assumed focal portion) of the line was in the lower half of the card. Finally, simply dividing the card itself, as in Cards 11–13, evoked the response of upside down when the line of division was in the lower portion. Card 14 appears to be an exception to this generalization, and constitutes the only instance (in this group of cards) in which the predicted choice was not made.

Fig. 3 shows the responses to the cards that were designed to explore the effects of differences in size and intensity in evoking preferences, or, in view of the results just described, in determining the focal portion. Black appeared to be more focal than white when the areas were equal (Card 5), even in the absence of a white background (cf. Fig. 1, Form 11). The preference for black on top was reversed, however, by the presence of dots in the white portion; that is, the dots were figural in contrast either to solid black (Card 6) or to solid white (Card 7). The effect of intensity was investigated further with the card consisting of three shades

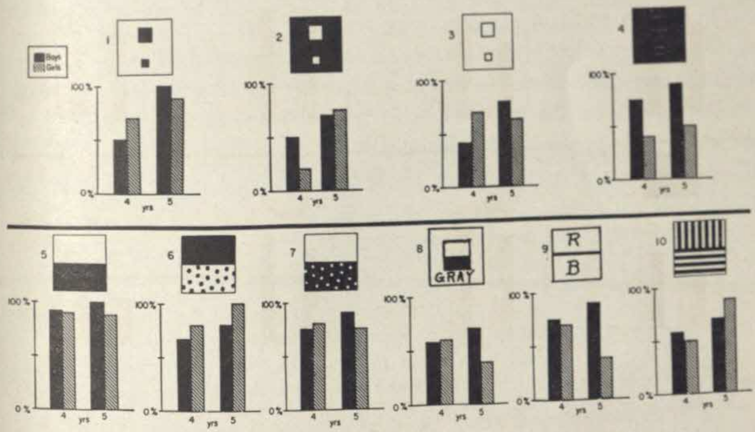


FIG. 3. PERCENTAGE OF BOYS AND GIRLS IN EACH AGE-GROUP CHOOSING CARD AS UPSIDE DOWN IN ORIENTATION SHOWN

Each card was 4 in. on a side and did not have the black border shown here. Card 8 consisted of three shades of gray, the background being intermediate in brightness between the dark and light portions of the figure. Card 9 consisted of a bright orange-red, and a dark royal blue.

of gray (Card 8); the responses were not different from chance. Preliminary investigation of responses to color (Card 9) indicated that 70% of the group considered the card upside down when the dark royal blue was at the bottom ($p < 0.02$); this preference could be due to difference either in color or in intensity.

The effect of the size of a figure in determining focal quality was investigated in relation to reversal of brightness between figure and ground, and in relation to the presentation of the figure in solid or outline form. The smaller figure consistently appeared to be the focal portion for the 5-yr.-old group, although the responses were significantly different from chance ($p < 0.02$) only in the case of the solid black square on the white ground. Consistent preferences did not appear in the 4-yr.-old group, and

the choices of the 4- and 5-yr.-old groups were significantly different ($p < 0.05$) for both cards with solid squares (Fig. 3, Cards 1 and 2). Reversal of the brightness-relations between figure and ground did not change the responses to the outline figures, but it did exert a significant effect with the solid figures. That is, 23 *S*s made the same response to Cards 1 and 2 (Fig. 3). Thirteen *S*s considered the card with the large square on top upside down when the background was white but right side up when the background was black, whereas only 4 *S*s showed the converse change. Evaluation of the change in response to different cards by the same individuals yields a chi-square for correlated proportions of 4.76 ($p < 0.05$).

The only significant sex-difference found in this experiment was in response to the outline white squares on the black ground (Fig. 3, Card 4), with the boys, in contrast to the girls, showing a strong preference for the smaller square on the top. Since this type of response (to the cards with the squares) appeared more frequently in the children of 5 yr. than in those of 4 yr., the boys' response could be interpreted as more mature than that of the girls.

The results of Experiment II confirm and extend the observations reported previously. Again, young children showed marked preferences for a particular orientation of nonrealistic forms, and, furthermore, these preferences appeared to be related to the position of the focal part of the figure. The reader may inquire whether these preferences can be obtained with other methods. In work not reported here, the cards have been shown singly and *S* asked whether each card was upside down or right side up; in other instances, *S* has been given the pile of cards, in random order and orientation, and has been required to place each card in the right side up position. For most figures, the preferences remained the same and as strong, regardless of the method of presentation used. The data of Experiment II will be discussed with respect to three findings concerning the determinants of correct orientation of nonrealistic figures: (1) the position of the focal portion; (2) the role of the vertical; and (3) the effects of brightness and size in defining the focal part.

(1) It was clear that, when the focal portion could be defined in an unambiguous way, the card was considered upside down if this focal portion was in the bottom part of the card. When a single homogeneous figure, such as a dot or a line, was placed in the lower portion of the page, the card was considered to be upside down (Fig. 2, Cards 1–4). When a simple figure was interrupted by a gap or a line, producing a non-homogeneous figure, the card was considered upside down when the interrupted portion was in the bottom half of the card (Fig. 2, Cards 7–10). When the card itself was divided into unequal portions of black and white, the card was considered upside down when the line of demarcation was in the lower portion of the card (Fig. 2, Cards 11–13). An apparent exception to this general finding was the absence of consistent responses to Card 14 in Fig. 2. In this instance, it is possible that some of the children were responding to the position of the figure on the card, in which case it would be upside down when the bottom part of the *card* contained the figure, whereas other children were responding to the orientation

of the figure itself, in which case it would be upside down when the bottom part of the *figure* contained the focal portion (cf. Fig. 1, Form 9).

(2) The vertical axis appears to play a role in determining correct orientation that is separable, at least in part, from the position of the focal portion. The children showed a strong preference for a centered homogeneous line in the vertical as opposed to the horizontal orientation. Incidental observations suggest that the preference for verticalization may sometimes override the tendency to place a homogeneous figure in the upper space of the card. For example, with the pair of cards consisting of a horizontal line in the upper portion of one card and a horizontal line in the lower portion of the other card (Fig. 2, Card 3), one *S* spontaneously called them both wrong. (When asked which one was more wrong, he pointed to the one with the horizontal line at the bottom.) Further, the *Ss* who were presented with the cards one at a time and asked to put them right side up often turned these cards in such a way that the lines were in the vertical position. Some of the children also turned the cards that were divided into black and white portions in such a way that the line of juxtaposition of the black and white was vertical rather than horizontal (Fig. 2, Cards 11 and 12).[3]

(3) A subsidiary purpose of the second experiment was to explore the effects of relative brightness and size in determining the focal aspect of a figure. If the young child prefers the focal part in the upper portion of the figure (as indicated by the responses to the cards in Fig. 2), then the part preferred on top can be used as a measure of the portion of the figure perceived as focal. Other factors being equal, it appeared that black was more focal than white (Fig. 3, Card 5), and there was a tendency for darker portions in general to be more focal than lighter portions (Fig. 3, Cards 8 and 9). When a large and a small figure were opposed, the small figure was the more focal one in the 5-yr.-old group, but the 4-yr.-old children were equally divided in their choices (Fig. 2, Cards 1–4). Similar results have been reported for adults in the perception of reversible figures, where there is some dominance of black over white, and of small areas over large areas, in determining the figure seen.[4]

INTERPRETATION

It is clear that the young child is not insensitive to the spatial position of forms; on the contrary, he has preferences for the orientation of nonrealistic forms that have been hitherto unsuspected by his adult examiners. It will be proposed here that these preferences derive from some characteristics of form-perception in the young child, and that perhaps form-perception in the adult is not entirely free of similar influences.

Development of form-perception. There is general disagreement in the literature as to whether one can appropriately talk of the development of

[3] There were two kinds of response in the single-card situation to the cards that were simply divided into black and white portions—the line of demarcation was verticalized, or was placed in the upper part of the card. Since only a small number of children were tested in the single-card condition, it is difficult to account for the two types of response, but there was a tendency for the younger children to put the line of demarcation on top and for the older children to verticalize it.

[4] H. Goldhamer, The influence of area, position, and brightness in the visual perception of a reversible configuration, this JOURNAL, 46, 1934, 189-206.

form-perception. The Gestalt view is that the perception of simple forms does not show any genetic development, whereas the opposite position has been taken by such diverse writers as Hebb[5] and Piaget.[6] The present writer takes the latter position and leans in particular on the framework provided by Hebb for the development of the perception of form.

Let us first assume that the process underlying form-perception in the young child involves a sequential consideration of the figure, or scanning of its various parts. Let us further assume that certain portions of a figure catch the attention, or 'draw the eye,' more readily than do other portions. The scanning of a figure would, then, begin with such a focal portion.[7] The results of the experiments just reported indicate that the young child prefers this focal part to be at the top of the figure, and he prefers the main axis of the figure to be vertical. These two preferences both could be manifestations of a tendency to scan a figure in a downward direction. That is, it is suggested that scanning of a form does not proceed in a random order from one focal part to another, but tends to start at the top and continue downward.

If there is a preference for scanning in a top-to-bottom direction (in the developing stage of form-perception), then the position of the figure can conform to, or be in conflict with, this tendency. For example, the 'V,' with the angle at the bottom, would contrast with the preferred sequence of scanning, since attention would be drawn first to the angle in the bottom portion of the figure in contrast to the preferred tendency to start at the top and continue downward. When the 'V' is oriented with the angle on top, attention again would be drawn to the angle first, and the lines of the figure then would be consistent with the tendency to scan downward. From this point of view, a geometric form is considered right side up when the position of the form on the card conforms to, or reinforces, the preferred sequence of scanning,[8] and the form is upside down when it is placed in a position that opposes the preference.

---

[5] D. O. Hebb, *The Organization of Behavior,* 1949, 17-106.

[6] Jean Piaget, *The Psychology of Intelligence,* 1947, 56-66.

[7] The reader may object that the form must be 'perceived' as a whole before the focal point can be 'identified' and serve as the starting point of a scanning process. The sense in which the term 'focal' is used here does not require that the whole be 'perceived' first, but it does require that the whole figure fall on the retina regardless of which part of the figure is fixated first. This condition is fulfilled, as can be readily seen from the size of the card and its distance from the eye. That one portion of this retinal pattern may be more effective than another in producing an eye-movement, or 'drawing the attention,' is a reasonable assumption and hardly an original one (see discussion by Hebb, *op. cit.,* 82).

Although the objection that the focal aspect can be perceived only secondarily has been raised predominantly by Gestalt psychologists, the objection is more in the spirit of Helmholtz' 'unconscious inference' than of Gestalt field-effects. In the Gestalt view, the focal aspect of a configuration would be given as a primary quality resulting from the simultaneous interaction of the various parts of the field. The position taken here is entirely consistent with the assumption of field-effects, but departs from the Gestalt view in that form-perception is not assumed to be a single, unitary process.

[8] 'Scanning' could refer to overt eye-movements or to an internal motor process, *i.e.* the motor facilitation of a cell-assembly suggested by Hebb. It has been found that eye-movements are not necessary to elicit preferences for orientation (Lila Ghent, Lilly Bernstein, and A. M. Goldweber, Preferences for orientation of form under varying conditions, *Percept. mot. Skills,* 11, 1960, 46). 'Scanning' is used here to

This interpretation of the young child's preferences for orientation implies that the child makes a judgment of orientation radically different from the usual one. It generally is assumed that the judgment refers to a comparison of the relation between a given form and its framework with the relation between another instance of the form and its framework. In contrast, it is proposed here that the perception of orientation in the young child is a judgment of whether or not the form is oriented in such a way as to conform with factors that operate in the actual perception of the form.

The possibility that different types of judgment are made by the child and the adult may explain why earlier workers in this area have concluded that the spatial orientation of a form is irrelevant to the young child. The observations that have been made indicate that the young child does not compare, or does not match, the relations between forms and their environments.[9] It is appropriate to conclude from such data that the orientation of a form, in the adult sense, is not perceived by the young child, but it is inappropriate to conclude that the orientation of a form is without influence on the child's perception.[10] The child's lack of response to orientation in the adult sense appears to be due to an inability to make the complex comparison of the relations required for the judgment. An alternative type of judgment has been proposed to account for the child's sensitivity to orientation, as indicated by his preferences for orientation.

It has been stressed that the child's judgment of orientation does not relate the position of a form to its framework. Nevertheless, the 'top' of the figure and the 'downward' direction of scanning must be defined with respect to a framework. In principle, these terms could be defined either with respect to an environmental framework or to a retinal framework; the argument developed here, however, would seem to require that the framework be a retinal one. In a separate study, Ghent, Bernstein, and Goldweber investigated the determinants of the phenomenal upright in preschool children by asking the children to judge the orientation of realistic and geometric figures while they were standing with their heads between their legs.[11] It was found that orientation was judged with respect to the retinal, and not the environmental, framework.

The interpretation offered for the preferences for orientation suggests that the judgment of the orientation of a figure is not separable, for the young child, from the perception of its shape. That is, a figure is called right side up if its position facilitates the tendency to scan in a top to bottom direction and hence facilitates the perception of the form. The burden

---

refer to an internal motor process facilitating perceptual activity in a particular sequence; this sequential organization conceivably could derive from overt eye-movements made in an early stage of development.

[9] Charlotte Rice, The orientation of plane figures as a factor in their perception by children, *Child Devel.*, 1, 1930, 111-43; Helen Davidson, A study of reversals in young children, *J. genet. Psychol.*, 45, 1934, 452-65; Davidson, A study of the confusing letters, B, D, P, and Q, *J. genet. Psychol.*, 47, 1935, 458-68.

[10] For example, it has often been claimed that young children recognize realistic figures equally well in any orientation. A direct test of this assumption has indicated that young children are particularly dependent on the right side up orientation of realistic figures for their recognition (Ghent, *op. cit.*, 249.)

[11] Ghent, Bernstein, and Goldweber, *op. cit.*, 46.

of proof of this argument rests on a demonstration that orientation does affect the recognition of form, and evidence in support of this position has been obtained. Ghent and Bernstein found that preschool children recognized geometric forms presented in the right side up position more frequently than forms presented in the upside down position.[12]

Changes with age. It has been suggested that the basis for the judgment of orientation is different in the preschool than in the older *S*s, and the ready acceptance by the younger *S*s of the problem of judging the orientation of non realistic forms did, indeed, contrast with the hesitant and questioning attitude of the older. The difficulty with this simple formulation is that the older *S*s were, nevertheless, able to make consistent judgments of the orientation of the nonrealistic figures, as discussed in Experiment 1; furthermore, similar judgments of (or preferences for) orientation of nonrealistic figures have been found in adults.[13] The difference between younger and older *S*s may be described by saying that the younger *S*s have only one basis for the judgment of orientation, whereas the older *S*s have two bases available; when the usual conceptual judgment cannot be made, the other basis for judgment comes into play. Although this formulation allows for the persistence of preferences in the older *S*s, it does not account for the fact that the pattern of preference in the older *S*s differs in at least two respects from that of the younger *S*s. It will be recalled that the number of cards evoking preferences in the older *S*s was less than in the younger *S*s, and that the older *S*s showed a few preferences that were significantly different from those of the younger *S*s.

The finding that children of school age and adults showed fewer preference for orientation than preschool children is difficult to interpret. To some extent, the older *S* may look for meaning in the nonrealistic representations in an attempt to make the judgment of orientation on the usual basis; this tendency, acting on different figures for different *S*s, would increase the probability of random responses appearing for the group. It is more likely, however, that the decrease in number of cards evoking preferences in the older *S*s reflects some more general change in perceptual processes with age. The most obvious possibility is that sequential consideration of a form plays a diminishing role in form-perception in the older *S*s; perhaps the definition of the focal part of a form does itself change with age.

The possibility that the focal aspect of a form may change is suggested, in part, by the responses to the figures with the gap (Fig. 1, Forms 7 and 8). The preschool children considered these figures upside down when the gap was on top, whereas

---

[12] Lila Ghent and Lilly Bernstein, Influence of the orientation of geometric forms on their recognition by children, *Percept. mot. Skills*, 12, 1961, 95-101.

[13] Significant preferences for the orientation of some nonrealistic figures have been found in two groups of college students, one group tested with the cards of Experiment I and the other group tested with those of Experiment II (Unpublished study).

the older children (and the group of adults referred to in Footnote 13) considered the figure upside down when the gap was on the bottom. This reversal of preference can most readily be interpreted by assuming that the line was focal for the younger *Ss* and the gap for the older *Ss*.[14] It will be recalled also that the response to the figures with the large and small squares (Fig. 3, Cards 1–4) showed some changes with age, suggesting that the small square might be considered focal for the 5-yr.-old *Ss* and the large square for the 4-yr.-old *Ss*. The possibility that the stimulus-characteristics defining 'focal' change with age in some instances raises difficulties for the definition of the focal part of a form, and there is risk of any definition becoming merely a tautology. To the extent, however, that further work can define the stimulus-characteristics preferred on top at different ages, the problems associated with the term 'focal' will be correspondingly diminished.

Although there is relatively little that can be said about developmental changes in the perception of form and its orientation after the preschool age, the existence of a type of judgment of orientation not heretofore considered has been suggested here. It seems reasonable to assume that the judgment of orientation of nonrealistic forms in the older *Ss* is closely related to the judgment of the preschool *Ss*. If the adult has a basis for judging the orientation of nonrealistic forms, and if this judgment indicates a tendency to scan form in a top-to-bottom direction, then one would expect these characteristics to manifest themselves in material other than that described here.

There are, in fact, a number of observations indicating asymmetry in the vertical axis in the perception of adults. Arnheim has reported that adults judge the orientation of abstract art with significant consistency—nonrealistic pictures appeared 'right' in one orientation and 'wrong' when rotated 90 or 180°.[15] Furthermore, all written languages of which the present writer is aware are read in a top-to-bottom direction rather than from the bottom to the top, although all other variations in direction seem to be present. Perhaps the asymmetry in the vertical axis most relevant to the results described here is that the perceptual center appears to be above the true center, as indicated by a constant error in the bisection of the vertical line,[16] and by the assumption by designers of advertising layouts that the 'optical center' is above (and to the left) of the midpoint of the page.[17] Presumably the same phenomenon accounts for the elementary rule in industrial design that a figure to be put in a frame, such as a picture on a mat or print on a page, must not be centered in the frame, since the figure would then appear to be below the center.[18] The apparent preference of adults for the placement of a figure above the true middle of the background might be described as an attenuated form of the preference shown by the young child for the placement of a figure in the upper portion of a card (Fig. 2, Cards 1–4).

[14] The Witte-König fusion-effect described by Helson and Wilkinson also indicates that a gap has a focal or dominant quality in the perception of adults (Harry Helson and A. E. Wilkinson, A study of the Witte-König paradoxical fusion-effect, this JOURNAL, 71, 1958, 316-20).

[15] Rudolph Arnheim, *Art and Visual Perception*, 1954, 18, 378.

[16] M. J. Delboeuf, Note sur certaines illusions d'optique, *Bull. de l'Académie Royale des Sciences, des Lettres, et des Beaux-Arts de Belgique*, 2me Série, 19, 1865, 195-216.

[17] Stephen Baker, *Advertising Layout and Art Direction*, 1959, 40.

[18] Harold Van Doren, *Industrial Design*, 2nd. ed., 1954, 151-52.

It has been proposed here that the sequential consideration of form in a downward direction plays a role in perception of young children and, to a lesser degree, in the perception of adults. It has already been suggested that sequential organization in a particular order could derive from movement of the eyes in scanning a form during the development of form-perception. The recent work of Ivo Kohler[19] and of Held and Hein,[20] in line with the theory of Von Holst,[21] indicates that organization of the perceptual world is strikingly affected by self-initiated movements of the organism; the role of such movements may be manifest primarily in the early stages of development (infancy) or of reorganization (*e.g.* after prismatic distortion).

SUMMARY

Investigation of the response of preschool children to pictures varying in orientation revealed that young children show remarkably consistent preferences for a particular orientation of nonrealistic figures as right side up. Such preferences are surprising, since they are clearly not learned from adults in the culture, and they seem to contradict the customary belief that young children are particularly unresponsive to the spatial orientation of visual forms.

Analysis of the stimulus-configurations eliciting preferences indicated that the young child prefers the focal part of the figure to be in the upper portion of the figure or the card; some data suggested, in addition, that the child prefers the main axis of the form to be in the vertical orientation. These preferences have been interpreted as indicating that, in the development of form-perception, scanning of form proceeds in a downward direction.

When the position of the focal part is such as to conflict with this sequential consideration of the form in a downward direction, the child considers the form to be upside down. Judgment of the orientation of a form is usually considered to be based on the relation of the position of a particular figure to the position of other figures in the environment, whereas it is suggested here that, for the young child, the judgment may be based on whether the position of the form facilitates, or conflicts with, movement-tendencies of the child.

---

[19] Ivo Kohler, Experiments with prolonged optical distortions, paper presented at the International Congress of Psychology, Montreal, Canada, 1954.

[20] Richard Held and A. V. Hein, Adaptation of disarranged hand-eye coördination contingent upon re-afferent stimulation. *Percept. mot. Skills*, 8, 1958, 87-90.

[21] E. von Holst, Relations between the central nervous system and the peripheral organs, *Brit. J. Animal Behav.*, 2, 1954, 89-94.

THE QUANTAL HYPOTHESIS AND THE THRESHOLD OF AUDIBILITY

By JOHN F. CORSO, The Pennsylvania State University

In a recent paper, Corso reviewed the neural quantum theory of sensory discrimination.[1] He pointed out that, among other factors, the resolution of the issues involved in this theory requires further work along two major lines: (a) "the development of a more satisfactory technique for statistically testing the goodness of fit of the quantal and phi-gamma hypotheses to a set of experimental data; and (b) the determination of the specific conditions under which rectilinear psychometric functions may be obtained."[2] The present study consists of two experiments related to these requirements. In these experiments, the neural quantum theory has been approached from a consideration of curves of absolute sensitivity, rather than from psychometric functions of differential sensitivity, as has been the case in most previous studies.

The first attempt to obtain evidence on the quantal nature of sensory functions from absolute thresholds was that of Békésy in 1936.[3] For a single *O*, minimum-audible-pressures (*MAP*) for pure tones were determined from about 2 to 50~ by alternately increasing frequency and decreasing intensity throughout the frequency-range tested. When the threshold curve was plotted with sound pressure in db. above 1 dyne/sq. cm. as a function of frequency, the curve showed a series of 'steps' which occurred at fairly regular intervals between 4 and 50~, with the most prominent step at 18~. These 'steps' were taken as an indication of the "quantum-like character of the nerve processes" involved in the determination of the absolute threshold.

In 1944, DeCillis used a quasi-quantal procedure to study the absolute threshold for tactual movement.[4] Three *O*s were required to indicate the presence or absence of movement as a fine column of air (at a pressure of 35 lbs./sq. in.) was moved over a cutaneous area on the fingertip, arm, or leg. The air-column was presented

* Received for publication January 19, 1960. This research was supported by the National Science Foundation. Experiment I was performed under Grant NSF-G1285; Experiment II, under Grant NSF-G4423. Acknowledgment is made to Alexander Cohen and Moira Fitzpatrick LaMay for their assistance in collecting and analyzing the data, and to Professor Walter L. Baker, for providing the special instrumentation used in the study.

[1] J. F. Corso, The neural quantum theory of sensory discrimination, *Psychol. Bull.*, 53, 1956, 371-393.

[2] *Idem*, 392.

[3] Georg von Békésy, Ueber die horschwelle und fuhlgrenze langsamer sinusformiger Luftdruckschwankungen, *Ann. d. Physik.*, 26, 1936, 554-566.

[4] O. E. DeCillis, Absolute thresholds for the perception of tactual movement, *Arch. Psychol.*, 294, 1944, 1-52.

at a point on the skin for 0.10 sec., traveled across the skin for a given distance at a rate of 143 mm./sec., and then remained stationary for at least 0.10 sec. Within the limitations indicated by Corso, the frequency of positive responses for a limited number of cases was found to be a linear function of the amplitude of stimulus-movement.[5] For these functions, however, the second prediction of the quantum theory did not hold, *i.e.* there was not an integral relation between the smallest stimulus-value at which 100% response occurred and the largest stimulus-value at which no response occurred. It was concluded that "apparently the integral relation is not to be expected in studies of absolute sensitivity."[6]

Except for these two studies, the theory of the neural quantum of sensory discrimination has not been approached from a consideration of absolute thresholds. It would appear, however, that if the concepts embodied in the theory are valid, evidence for or against the theory should be obtainable from either absolute or differential thresholds.

Since the results of earlier studies dealing with differential sensitivity have not resolved the issues inherent in the quantal and phi-gamma hypotheses, the development of an alternative approach is presented in an attempt to establish more definitive findings.

EXPERIMENT I

The present study was designed to replicate and to extend the initial work of Békésy who reported that the cuve of audibility for pure tones of low frequency revealed a quantal process. Several *O*s, rather than one, were used and threshold-measurements were made for 5~ to approximately 200~ to determine whether the upper limit at which the 'quantal steps' could no longer be detected extended beyond the original estimate of 50~. It was thought that the method used in establishing the audibility-curve would yield 'steps' beyond 50~ to the upper frequency-limit to be tested, *i.e.* 200~.

Procedure. (1) Observers. The *O*s, 15 in number (5 men and 10 women), were obtained from an undergraduate course in educational psychology. They ranged in age from 19 to 22 yr., with the exception of one man who was 27 yr. old. All the *O*s had normal hearing.[7] To maintain a high level of motivation throughout the experiment, the *O*s were paid for participation in the study.

(2) Apparatus. The apparatus used in certain preliminary practice-periods and in

---

[5] Corso, *op. cit.*, 387.

[6] Corso, *op. cit.*, 558.

[7] Audiograms obtained from the *O*s showed that none deviated from the standards set by the American Medical Association by more than 5 db. at any point between 250 and 2000~. For the reference-curve, see Anon., Minimum requirements for acceptable pure-tone audiometers for diagnostic purposes, *J. Amer. Med. Assoc.*, 146, 1951, 255-257. Otological examinations were also made on all the *O*s by Drs. H. R. Glenn and E. S. Krug of the University Health Service, to whom acknowledgment is extended for their coöperation. With the exception of one *O* who showed some evidence of injection of both tympanic membranes and the minor presence of calcareous deposits, there were no other observable defects.

the absolute-threshold determinations consisted of four major units: (1) a low-frequency oscillator—Hewlett-Packard, Model 200 DC; (2) a decade-attenuator—General Radio, Type 1450-TB; (3) a power-amplifier; and (4) a specially-designed receiver. The receiver was coupled to the *O*'s ear by means of a small brass tube which was sealed in the external auditory meatus with a mixture of plastic modeling clay and vaseline.

The frequency-response, harmonic distortion, and sound-pressure output of the low-frequency system were determined through calibration measurements with the output of the receiver directed into a National Bureau of Standards (NBS) Coupler 9-A.[8] A detailed description of the system and its technical characteristics may be found in an earlier publication.[9]

(3) Preliminary practice. In preparation for the main testing session, each *O* was given a preliminary series of practice-tests. Six tests were used. (1) The Seashore pitch-tests (Forms A and B) were presented by means of a Grason-Stadler speech audiometer, Model 1160-A. (2) An audiometric test was given on both right and left ears by means of a Beltone audiometer, Model 10-A, modified for Permoflux PDR-8 earphones. In this test, the method of limits was used with the stimulus-intensity varied in 2-db. steps, three tonal pulses at each step. (3) The right ear only was tested by means of Békésy-type audiometer, Grason-Stadler, Model E-254. In this test, the frequency range from 250 to 8000~ was presented in both an ascending and descending order with an attenuation rate of 0.7 db. per sec. and a total testing-time of 11 min. for each order. (4) A re-test was given on the Beltone audiometer for the right ear only. (5) Threshold-measurements were made on the right ear only at 5, 25, 50, 75, 100, 125, 150, 175, and 200~ by means of the low-frequency system. The method of limits was used with a 5-db. step between successive stimulus-intensities, one pulse per step. (6) Threshold-measurements were made as in (5) with the exception that a 2-db. step was used, three pulses per step. Each of these pre-tests was administered on a different day; the total practice-time was approximately 6 hr.

(4) Test-session. After each *O* had completed the preliminary practice, measurements were made of minimal audible-pressures (for the right ear only) from 5 to 200~ in both ascending and descending orders. In the ascending order, the signal was introduced at 5~ at an intensity-level well above threshold.[10] *E* then decreased the intensity in 1-db. step (with *O* making a judgment at each setting) until *O* failed to respond; at this intensity-level, *E* slowly and continuously increased the frequency by means of an automatic drive-mechanism which turned the shaft of the frequency-dial on the Hewlett-Packard oscillator at a rate of 0.75 r.p.m. When *O* pressed the response-button indicating that he heard the tone, the drive-mechanism was immediately stopped. At this frequency, the intensity was once again decreased stepwise until *O* failed to respond; thereupon the frequency was again increased until an auditory response was obtained. This procedure was continued until the test-frequency reached 200~.

[8] American Standards Association, Inc., *American Standard Method for the Coupler Calibration of Earphones*, 1949.

[9] Corso, Absolute thresholds for tones of low frequency, this JOURNAL, 71, 1958, 367-374.

[10] The testing procedure described for the ascending sequence followed as closely as possible that reported by Békésy (*op. cit.*, 558). In that study, however, there was no descending sequence of frequencies.

In the descending sequence, the signal was introduced at 200~ at the last audible intensity-level reached in the ascending sequence. At this level, the signal was automatically *decreased* in *frequency* until *O* failed to hear the tone; at this frequency, *E increased* the *intensity-level* to the next to the last intensity-setting obtained in the ascending sequence. As this elicited a response, the frequency was once again continuously decreased until *O* failed to hear the tone. This procedure was systematically repeated until the frequency reached 5~.

All measurements for both testing sequences were completed in a 1-hr. session for each *O*, with a short rest-period between successive sequences. These measurements, as well as those in preliminary practice, were made in an anechoic chamber provided with a two-way voice-communication system and a closed-circuit television system for visual monitoring of *O*'s activities during the experimental sessions.[11]

(5) Treatment of data. The experimental values obtained for each *O* were graphically plotted with frequency and intensity arbitrarily assigned to the abscissa and ordinate, respectively. Since the ascending and descending curves were superimposed, this provided a total of 15 sets of curves, one set for each *O*. The threshold-curves were then analyzed to determine: (1) whether the obtained, ascending curves approximated the 'quantal' threshold-curve of Békésy,[12] and (2) whether the ascending and descending curves for each *O* showed similarities with respect to the location of the 'steps' in frequency.

To determine the *maximal* extent of agreement between the data of the present study and those of Békésy, product-moment correlation-coefficients *(r)* were computed between the values of Békésy's curve and those from the corresponding portion of the ascending curve for each *O*. Each correlation coefficient was based upon the frequency-location of the 'quantal steps' obtained by Békésy (*i.e.* 4.5, 6, 7.5, 9, 11, 14, 18, 22, 28, 32, and 38~) and the *closest* corresponding frequency in the ascending curve for each *O*. Values in either curve for which no corresponding 'steps' could be determined were omitted in the computation of the correlation-coefficients.

Results. The correlation-coefficients obtained for the 15 *O*s are presented in Table I. Notice that the values range from 0.983 to 0.999. Thus, most cases show excellent agreement with Békésy's data, considering only the *frequency-location* of the 'steps' and not their intensive magnitudes.

A segment of the ascending curve for *O* No. 6, with Békésy's curve superimposed, is shown in Fig. 1.[13] Note the close agreement in both the frequency-location of the 'quantal steps' and their relative magnitudes. For ease of comparison, Békésy's curve was shifted downward by 12 db. to account for the difference in the auditory sensitivity of the two *O*s.[14]

[11] R. L. Berger and E. Ackerman, The Penn-State anechoic chamber, *Noise Control*, 2, 1956, 16-21.

[12] Békésy, *op. cit.*, 558.

[13] As can be seen in Table I, other curves yielded higher correlation-coefficients than that of *O* No. 6; however, this curve was selected since it was based upon the greatest number of experimental points ($N = 11$).

[14] Békésy, *op. cit.*, 558.

The entire form of the function from 5 to 200~ obtained for a typical *O* under the ascending and descending conditions of the present study is shown in Fig. 2. The data are for *O* No. 10. The solid line in Fig. 2 shows the function for the ascending condition; the broken line, for the descending condition. As hypothesized, discontinuities in the curves can be observed throughout the frequency-range tested. Similar discontinuities were found for all other *O*s.

To evaluate the reliability of measuring the location in frequency of the discontinuities in the ascending and descending threshold-curves, a second set of correlation-coefficients was computed. For each *O*, the frequency-locations of the 'steps' in each experimental curve were paired on the basis of the sound-pressure levels at which the ascending and descending determinations were made. The product-moment correlation-coefficients computed for these paired frequency-values for each *O* are presented in Table II. All correlation-values are exceedingly high, ranging from 0.971 to 0.998. It appears, therefore, that there is considerable agreement in the location of the 'steps' in frequency as determined from the ascending and descending conditions of the present study.

TABLE 1

PRODUCT-MOMENT CORRELATION-COEFFICIENTS COMPUTED BETWEEN THE FREQUENCY-LOCATION OF THE DISCONTINUITIES IN BÉKÉSY'S THRESHOLD-CURVE AND THOSE OBTAINED IN THE PRESENT STUDY

O	*N**	*r*	*O*	*N*	*r*	*O*	*N*	*r*
1	7	.999	6	11	.991	11	10	.994
2	8	.997	7	6	.997	12	10	.996
3	10	.997	8	10	.984	13	10	.998
4	8	.983	9	3	.994	14	6	.997
5	6	.999	10	10	.997	15	8	.998

* Since there were only 11 discontinuities or 'steps' in frequency in Békésy's threshold-curve from 4.5 to 38~, the maximal *N* possible for the computation of *r* was 11 in each case.

Discussion. The results of this study as presented in Table I indicate that in general the ascending experimental curves are in close agreement with the form of the audibility-function obtained by Békésy under similar testing-conditions.[15] One major difference, however, exists. As indicated in Fig. 2, the experimental curves of the present study show discontinuities up to 200~, as contrasted to those of the earlier study in which the discontinuities disappeared at about 50~. Since the curves of all *O*s con-

[15] *Op. cit.*, 558.

sistently showed this effect, the question arises as to whether these 'steps' are really indications of quantal processes, as previously suggested, or merely artifacts introduced by the experimental methodology.

A careful reconsideration of the methodology indicates three critical features: (1) *O* was presented at each stimulus-magnitude, *i.e.* a particular combination of frequency and sound-pressure level, with only *one tone-pulse* to which he responded in a 'yes' or 'no' manner. Following this, another stimulus-magnitude was presented. This technique assumes complete reliability of *O*'s response; furthermore, it forces the response into

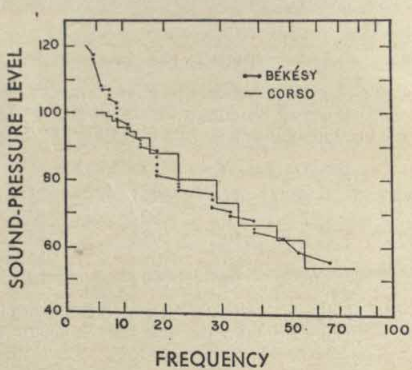


FIG. 1. COMPARISON OF A SEGMENT OF THE ASCENDING EXPERIMENTAL FUNCTION FOR *O* No. 6
(To equate for differences in sensitivity, Békésy's function has been displaced downward by 12 db.)

discrete all-or-none categories. (2) When, at a given frequency, the sound-pressure of the stimulus reached a value below threshold, *E* thereupon increased the frequency (holding sound-pressure constant) until the tone became audible. This procedure, regardless of the character of the underlying neural or quantal processes, must necessarily produce 'steps' in the threshold-curve provided two or more presentations of the (intensive stimulus are made at a given frequency. (3) The experimental points which determine the form of the 'audibility'-curve do not all have the same significance. Some points represent audible stimulus-values; others, non-audible values. Thus, a single function cannot meaningfully be plotted connecting all experimental points, as has been done in Figs. 1 and 2.

A solution to these difficulties is suggested in Fig. 3, where a portion of an 'audibility'-curve, based upon single judgments at each experimental point, is shown by the steplike function connecting a set of hypothetical values. The values of the steplike function, indicated by the dots, represent

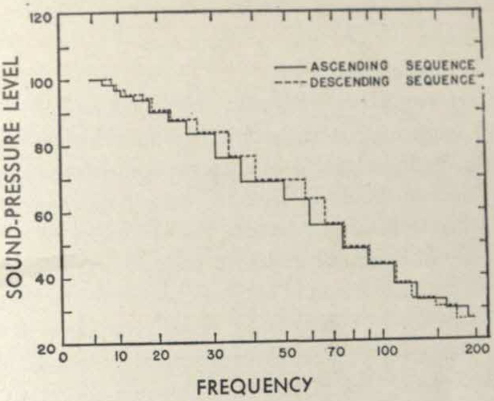


FIG. 2. COMPARISON OF EXPERIMENTAL FUNCTIONS OBTAINED UNDER ASCENDING AND DESCENDING CONDITIONS OF FREQUENCY-CHANGE
(Data are for *O* No. 10)

combinations of frequency and sound-pressure which are audible, whereas the open circles represent inaudible conditions. A more appropriate audibility-function can be obtained by connecting all experimental points representing the same probability of response, *e.g.* connecting all open circles or all points where a given sound-pressure is first reported as audible as

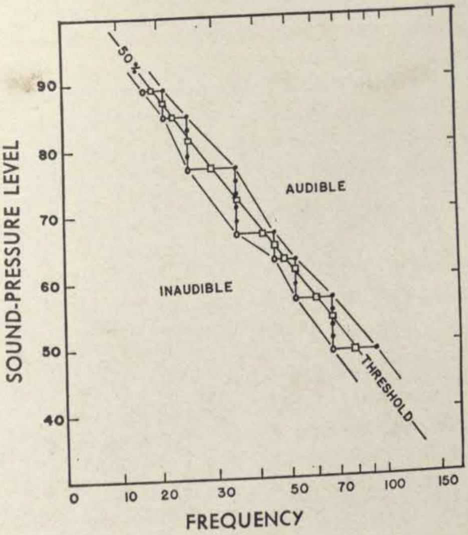


FIG. 3. THEORETICAL DERIVATION OF THE CLASSICAL AUDIBILITY-FUNCTION

frequency is increased. The two curves which represent these conditions are shown as broken lines in Fig. 3 and are used to bound the inaudible and audible areas, respectively.

This latter set of curves, however, while overcoming some of the methodological limitations, is still subject to the criticism that each curve is based upon a single judgment at each experimental point. Thus, some psychophysical method should be used to generate more judgments at each stimulus-magnitude to take into account the variability of *O*'s judgments. If this were done, each vertical and each horizontal segment of the 'steps' shown in Fig. 3 would correspond to a psychometric function. In Fig. 3, a solid line has been drawn connecting a series of squares representing the assumed 50%-points for this series of (quantal or phi-gamma) psychometric functions. Notice that in this case, although the 'steps' are unequal

TABLE II

PRODUCT-MOMENT CORRELATIONS COMPUTED BETWEEN THE FREQUENCY-LOCATIONS OF THE DISCONTINUITIES IN ASCENDING AND DESCENDING EXPERIMENTAL CURVES

O	*N**	*r*	*O*	*N**	*r*	*O*	*N**	*r*
1	15	.972	6	16	.993	11	17	.998
2	16	.996	7	15	.977	12	17	.993
3	21	.997	8	19	.997	13	15	.995
4	15	.996	9	10	.989	14	10	.989
5	13	.971	10	17	.997	15	16	.996

* *N* represents the number of paired frequency-values.

in both frequency and intensive magnitude, the discontinuities have disappeared and that a conventional audibility-function has been derived. Thus, from these considerations, it appears that the 'quantal steps' previously obtained in the auditory threshold-function and reproduced in the present study can be attributed to experimental methodology, rather than to underlying neurological or psychological processes.

The results shown in Table II, which are inherently subject to the same criticisms made for the data summarized in Table I, do nevertheless indicate that the experimental procedures produce reliable results. At a given sound-pressure level, the same frequency is 'barely audible,' whether the threshold is approached from a lower frequency in an ascending direction or from a higher frequency in a descending direction. A review of the 15 sets of threshold-curves (one set for each *O*) shows that the discrepancies in the frequency-location of the 'quantal steps' can be accounted for in terms of errors of habituation. In nearly every instance, the frequency-

location of the 'steps' was lower in the descending sequence than in the ascending sequence. This is typical of the type of error one expects to find when using the method of limits in measuring absolute thresholds and tends to support the notion that the 'quantal steps' are artifacts generated by the experimental procedure.

Experiment II

Experiment II was performed to permit a reëvaluation of the quantal hypothesis as applied to the absolute threshold of audibility, taking into account the methodological limitations suggested in Experiment I. Specifically, Experiment II was designed to determine the tenability of the neural quantum theory of sensory discrimination under conditions in which audibility-functions were established with *repeated* response-measures at each stimulus-value. Furthermore, a more exacting statistical analysis of the data was performed to compare the goodness of fit of a linear function to values independently predicted from the quantal hypothesis and the phi-gamma hypothesis.

Procedure: (1) Observers. The *O*s of this study were 10 volunteers (5 men and 5 women) from an undergraduate course in educational psychology. Their ages ranged from 19 to 23 yr. inclusive. All the *O*s were screened for normal hearing and otological disorders as indicated in Experiment 1. Also, *O*s were paid at the same rate for participating in the experiment. No *O* used in Experiment I was included in Experiment II.

(2) Preliminary practice. The low-frequency system, calibration-procedures, and testing-chamber were the same as in Experiment I. In Experiment II, however, the preliminary practice-period was modified to provide more opportunity for listening to low-frequency signals. The practice-tests which required about 5 hr., included Tests (b), (d), (e), and (f) of Experiment I and, in addition, a preliminary quantal test. The preliminary test was intended to establish for each *O* the size of the intensity-decrements and frequency-increments which would be used to obtain psychometric functions in the final test-period. This was accomplished by starting at 5~ with an intensity-setting which yielded approximately 100% response in 25 stimulus-presentations and then reducing this setting in discrete steps until 0% response was obtained. Likewise, at the intensity-setting at which 0% response was obtained, frequency was discretely increased in small steps until 100% response was obtained. This procedure was replicated at 10, 15, 20, 25, and 30~. On the basis of these data, an average value was obtained for each *O* for the intensity-change which corresponded to a decrement in correct response from 100 to 0%, and an average frequency-change was obtained which yielded an increment from 0 to 100% correct responses. From these average values, estimates were made of the step-magnitudes of intensity-change and frequency-change which should be used for each *O* in the final test-period such that each psychometric function would be based upon at least three values of the physical stimulus, omitting the 0% and 100% points of response.

(3) Quantal test-session. In the quantal test-sessions, the same general procedure was followed as described in Experiment I, with three major differences. (1) All tests were started at 5~ and terminated after *O* had reached a frequency of about 25~. This upper limit was adopted because collecting data up to this point usually required at least 1-hr. of testing-time, beyond which *O* became increasingly fatigued. (2) Only the ascending sequence of frequency-presentation was used. (3) At each experimental setting of frequency and intensity, a series of 25 pulses was presented. *O* responded to each audible signal by depressing a mercury switch which activated a Veeder-Root counter at *E*'s position. The number of auditory stimuli in a series was controlled by an arrangement of Hunter decade, interval-timers, Model 111-A, which were also used to maintain the stimulus-duration at 0.5 sec., with 2.5-sec. interval between successive stimuli. To minimize the possibility of extraneous signals entering the receiver and providing the *O* with irrelevant cues, the sinusoidal signal from the low-frequency system was passed through a narrow-band filter, Krohn-Hite model 330-M, preset at each test-frequency to yield minimal transient effects.

(4) Treatment of data. The data of the present study were treated in a four-step sequence: *(1)* For each *O*, the best-fitting linear and ogival psychometric functions were determined by the method of least-squares for all intensity- and frequency-conditions.[16] All functions were based on 3 to 17 empirical points, with judgments of 0 to 100% omitted. No weights were used in determining the curves of best fit. The effect of this step in the analysis of data was to transform the steplike 'threshold'-curve for each *O* into a series of psychometric functions in which intensity and frequency alternated as the independent variable.

(2) From the equations of the fitted functions, the predicted stimulus-values in intensity and in frequency corresponding to the 90% response-points were then computed. This yielded in each case a different prediction by the quantal hypothesis and the phi-gamma hypothesis. It should be pointed out that 90% of response represents approximately the point of maximal discrepancy between the two theoretical psychometric functions; the stimulus-values corresponding to the 50% response-points are the same for the two functions. By referring to Fig. 3, it can be seen that this procedure had the effect of shifting the squares in the figure in two ways: (a) from a 50% to a 90% criterion of response for all segments of the step-like function; and (b) in an amount corresponding to the predictions made by the quantal and phi-gamma hypotheses.

(3) The next step in the analysis was to derive for each *O* a set of best-fitting 'absolute threshold'-curves (sound-pressure level as a function of frequency) based upon the predicted stimulus-values corresponding to the 90% of response-points. Since there were two independent variables under study (frequency and intensity) and for each of these the 'true' psychometric function could be either linear or ogival, it was possible to obtain four different 'threshold'-curves for each *O*. These represented curves based upon the theoretical predictions in which (a) both frequency- and intensity-functions were assumed to be linear; (b) both frequency- and intensity-functions were assumed to be ogival; (c) the frequency-function was assumed to be linear and the intensity-function, ogival; and (d) the frequency-function was assumed to be ogival and the intensity-function linear.

---

[16] J. P. Guilford, *Psychometric Methods*, 1954, 63-66, 125-129.

For each of these four conditions for each *O*, the 'threshold'-curve of best fit was obtained by the method of least-squares for the predicted stimulus-values at the 90% response-points. Each curve was based upon four to nine predicted values. It was assumed that sound-pressure level at the 90% threshold was a linear function of the logarithm of frequency. This assumption appeared reasonable since only a very limited portion, *i.e.* about 25~ of the audibility function was tested in this experiment. Furthermore, previous research at low frequencies has indicated the existence of such a relationship.[17]

(4) The goodness of fit of each of the four 'threshold'-functions for each *O* was expressed by the index of correlation, $p_{y\ f(x)}$.[18] These indices for all *O*s were converted into z-measures computed from logarithmic transformations and were treated in a conventional analysis of variance (treatments-by-subjects design).[19] Tests of significance were performed to determine whether the goodness of fit of the '90% threshold'-curve was affected by the four different theoretical conditions.

TABLE III

INDEX OF CORRELATION FOR THE BEST-FITTING FUNCTION (90% AUDIBILITY) OBTAINED UNDER EACH OF FOUR THEORETICAL CONDITIONS

		Combination of hypotheses			
O	Frequency: Intensity:	quantal vs. quantal	phi-gamma vs. phi-gamma	quantal vs. phi-gamma	phi-gamma vs. quantal
1		−0.946	−0.971	−0.954	−0.963
2		−0.963	−0.972	−0.968	−0.967
3		−0.985	−0.989	−0.984	−0.986
4		−0.891	−0.908	−0.912	−0.884
5		−0.968	−0.976	−0.977	−0.966
6		−0.950	−0.946	−0.947	−0.949
7		−0.994	−0.982	−0.985	−0.959
8		−0.961	−0.977	−0.962	−0.976
9		−0.879	−0.875	−0.885	−0.869
10		−0.949	−0.832	−0.841	−0.945
	Mean*	−0.961	−0.960	−0.956	−0.957

* Computed from z-transformations of the tabled $p_{y \cdot f(x)}$-values and reconverted into the mean-values shown.

Results. Table III presents the index of correlation obtained for each of the 10 *O*s under each of the four theoretical conditions. These values range from −0.832 to −0.994; but the mean values for the four conditions range only from −0.956 to −0.961. Thus, there appears to be relatively little difference in the goodness of fit produced by the four theoretical conditions.

[17] Corso, *op. cit.*, this JOURNAL, 71, 1958, 372.
[18] Don Lewis, *Quantitative Methods in Psychology*, 1948, 81-82.
[19] W. J. Dixon and F. J. Massey, *Introduction to Statistical Analysis*, 1951, 164-166; E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 156-167.

Table IV presents a summary of the analysis of variance performed on the z-transformed measures of goodness of fit. Although there are significant differences among *O*s, the *F*-ratio (0.270 with 27 and 3 *df.*) for evaluating the differences among theoretical conditions is not statistically significant. The results of this analysis indicate that there is no difference in the quantal and phi-gamma hypotheses with respect to the prediction of the 90% audibility function; hence, both hypotheses are equally acceptable.

Discussion. The theory of the neural quantum and the classical phi-gamma hypothesis represent theoretical attempts to account for sensory discriminatory behavior. As previously indicated, most studies on the quantum theory have involved data obtained in differential threshold-measurements. For the most part, the previous studies have attempted to evaluate the quantal hypothesis on the basis of two criteria applied to a single psychometric function: (1) a linear relationship between the percentage

TABLE IV

SUMMARY OF ANALYSIS OF VARIANCE OF z-TRANSFORMED MEASURES OF GOODNESS OF FIT (INDEX OF CORRELATION)

Source of variation	Sum of squares	*df.*	*F*-ratio
Hypotheses (*H*)	0.031	3	0.270
Observers (*O*)	5.365	9	16.108*
$H \times O$	1.001	27	
Total	6.397	39	

* $p > 0.01$.

of stimulus-increments heard and the magnitude of the increments presented; and (2) a two-to-one ratio between the stimulus-values of the psychometric function at the 100% and 0% points of response.

In the present study, the evaluation of the quantal and phi-gamma hypotheses was based upon the index of correlation for the 90% audibility-curve best-fitted to a set of values predicted from several psychometric functions for a given *O*. In this evaluation, repeated measures (judgments) were obtained at each value of the stimulus. Thus, Experiment II avoided the methodological limitations pointed out in Experiment I. It also avoided the weaknesses of some earlier quantal studies in which the chi-square test was applied to differences between obtained and predicted values for single psychometric functions.[20] Despite a more rigorous ap-

[20] For a summary of this problem, see Corso, *op. cit., Psychol. Bull.*, 53, 1956, 381-388.

proach, the data of the present study have failed to yield unilateral support for either the quantal or phi-gamma hypothesis. As in most previous studies, both hypotheses remain tenable. It appears, therefore, that although two distinct theories of sensory discrimination have been formulated, appropriate experimental techniques have yet to be devised for the differential evaluation of the specific hypotheses under consideration.

SUMMARY

The present study consisted of two experiments designed to investigate the neural quantum theory of sensory discrimination through a consideration of the audibility-curve for tones of low frequency.

In Experiment 1, the general method was to determine the threshold of hearing in a continuous manner from 5 to 200~. For an ascending direction of frequency-change starting at 5~, a testing-procedure was used in which, alternately, frequency was increased and intensity decreased. For the descending direction starting at 200~, frequency was decreased and, alternately, intensity was increased. Fifteen selected *O*s were tested, each of whom showed: (a) negative findings upon otological examination and (b) no hearing loss greater than 5 db. in the range from 250~ to 2000~ as measured by a standard audiometric test.

After the *O*s were given 6 hr. of training on various tests, audibility-functions were obtained for right ears only from 5 to 200~. Stimulus-intensities were expressed in db. SPL *re* 0.0002 dyne/sq. cm. as measured in a National Bureau of Standards Coupler 9-A. In the analysis of data, product-moment correlation-coefficients were computed: (a) between the frequency-location of the 'steps' in the ascending audibility function for each *O* and the 'steps' obtained by an earlier investigator; and (b) between the frequency location of the 'steps' in the ascending and descending functions for each *O*. In both analyses, all obtained correlational values were above 0.97.

The interpretation of results suggested that the 'steps' in the audibility-function, rather than representing neural or quantal processes, could be accounted for in terms of the experimental methodology. The high intra-*O* correlational values in the ascending and descending trials were interpreted as an indication of the reliability of judgments obtained under the testing-conditions of the present study. It was concluded that 'steps' in an audibility-function derived from the procedures used in this experiment do not in themselves provide support for the neural quantum theory of sensory discrimination.

In Experiment II, the general procedure followed that indicated for Experiment I, with three exceptions: (1) the absolute threshold was determined from only 5 to about 25~ for each *O;* (2) only the ascending direction of frequency-change was used; and (3) a series of 25 presentations, rather than a single presentation, was made at each experimental value of the stimulus. Ten *O*s, all of whom met the same screening criteria as in Experiment I, served in Experiment II. After the *O*s were given about 5 hr. of training in a series of tests which included some of those administered in Experiment I, audibility-functions were obtained for right ears only from 5 to about 25~.

In the analysis of data, these audibility-functions were divided into segments to which rectilinear (quantal) and ogival (phi-gamma) psychometric functions for frequency and intensity were fitted by the method of least squares. The '90% response-points' predicted by these derived functions were in turn fitted by a linear function under four different theoretical conditions: (1) linear psychometric functions for frequency and intensity; (2) ogival psychometric functions for frequency and intensity; (3) linear psychometric functions for frequency, and ogival for intensity; and (4) ogival psychometric functions for frequency and linear for intensity.

The goodness of fit of each of the four linear functions derived from the data of each *O* was expressed by the index of correlation. An analysis of variance on the z-transformed correlational indexes showed no significant differences among theoretical conditions. Thus, the results of Experiment II indicate that both the quantal and phi-gamma hypotheses are equally acceptable in predicting the 90% audibility function for pure tones. It appears that although the phi-gamma hypothesis and the neural quantum theory of sensory discrimination postulate fundamental differences in discriminatory processes, adequate experimental techniques have not yet been devised which will yield data in support of one theory to the exclusion of the other.

MULTIPLE-CHOICE DECISION-BEHAVIOR WITH DUMMY CHOICES

By R. ALLEN GARDNER, Wellesley College

The consequences of a choice are commonly determined by critical events that follow the choice, and, more often than not, these critical events are to some degree uncertain at the time of choice. The present study was concerned with choices of this kind as they are made in a laboratory by a method generally attributed to Humphreys.[1]

In this task, the subject (S) attempts to anticipate a series of events. Each member of the series belongs to one of a set of categories, $A, B, C, \ldots, N$, and the categories are represented by suitable stimulus-objects which correspond to the set of responses that are available to S. If the series is unpredictable, then correct anticipations, or hits, will occur only by chance agreement between responses and events. If the relative frequencies of the categories, $\pi(A)$, $\pi(B)$, $\pi(C)$, $\ldots$, $\pi(N)$, are unequal, however, S can improve his performance up to a maximum by choosing the most frequent category on every trial.

The extent to which a series of choices will approach the optimal strategy is only partly dependent upon $\pi(A)$, the relative frequency of the most frequent category. For example, the number of categories has been varied with $\pi(A)$ held constant.[2] Under these conditions, $P(A)$, the rate of choice of the most frequent category, A, increased with the number of alternatives, $B, C, \ldots, N$, after a sufficient number of trials.[3] The procedure employed in these multiple-choice experiments required that one event occur on each trial, and that all of the categories available as choices be represented in the program of events. Thus, as the number of categories was increased, holding $\pi(A)$ constant, the average frequency of the minor categories decreased to meet the requirement $\pi(A) + \pi(B) + \pi(C) + \ldots + \pi(N) = 1.0$. Consequently, the results cannot be interpreted simply in terms of the increase in the number of choices made available to S. At least two other interpretations are

* Received for publication February 11, 1960. This research was carried out at the US Army Medical Research Laboratory, Fort Knox, Kentucky, as a subtask of Project No. 6-95-20-001. The author wishes to acknowledge the technical assistance of John Forsythe, Richard Kling, and Karl Verhoye as well as the advice and encouragement of Dr. George S. Harker.

[1] L. G. Humphreys, Acquisition and extinction of verbal expectations in a situation analogous to conditioning, *J. exp. Psychol.*, 25, 1939, 294-301.

[2] R. A. Gardner, Probability-learning with two and three choices, this JOURNAL, 70, 1957, 174-185; Multiple-choice decision-behavior, this JOURNAL, 71, 1958, 710-717; J. W. Cotton and Allan Rechtschaffen, Replication report: Two- and three-choice verbal conditioning phenomena, *J. exp. Psychol.*, 56, 1958, 96.

[3] With too few trials the effect does not always appear. Cf. M. H. Detambel, A test of a model for multiple-choice behavior, *J. exp. Psychol.*, 49, 1955, 97-104; E. D. Neimark, Effects of type of nonreinforcement and number of alternative responses in two verbal conditioning situations, *ibid.*, 52, 1956, 209-220.

possible. First, the rate of choice, $P(A)$, may depend not on the ratio of hits to misses obtained by choosing cateory A, but on this ratio relative to the average ratio of hits to misses obtained by choosing an alternative category. In this case, the multiple-choice procedure that was employed may have served merely to alter this relationship by lowering the ratio of hits or raising the ratio of misses that could be obtained by a choice of any one of the non-A categories. Secondly, a process of perceptual contrast may be operating in the programs, causing S to overestimate $\pi(A)$. That is, $\pi(A)$ may be perceived as more frequent in the context of a program containing more and more minor categories of lower and lower frequency.[4]

In the previous experiments, the multiple-choice effects need not have been a function of the increase in the number of choices in and of itself but rather they may have been a function of other variables confounded in the procedure used to increase the number of choices. With a modified procedure, the present study aimed at reproducing these multiple-choice effects in such a way that they could only be attributed to the number of choices. In the modified procedure, both the number and the relative frequencies of the categories in the program were held constant, while additional categories of choice were made available to S. These additional categories were introduced merely by adding to the number of response switches that S could operate. They will be referred to as 'dummy choices' since they never appeared in the program of events and no hits could be obtained by choosing them.

The objective of this study may be stated as follows. Hypothesis 1: After a sufficient number of trials, $P(A)$, the rate of choice of A, the most frequently presented category, varies directly with the number of categories of choice available to S, even when the number of categories of events presented in the program is held constant. That is, for a given program of critical events, $P(A)$ will be a direct function of the number of dummy choices. In addition, the data that were obtained are relevant to the hypothesis of asymptotic matching which may be stated in two parts as follows.

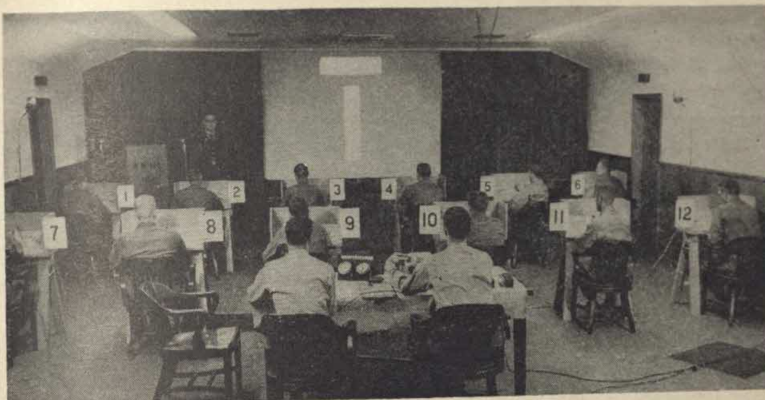
Hypothesis 2 a: $P(A)$ approaches an asymptote which is less than 1.0.

Hypothesis 2 b: The asymptote of $P(A)$ is determined by $\pi(A)$ in such a way that $P(A) \leq \pi(A)$.

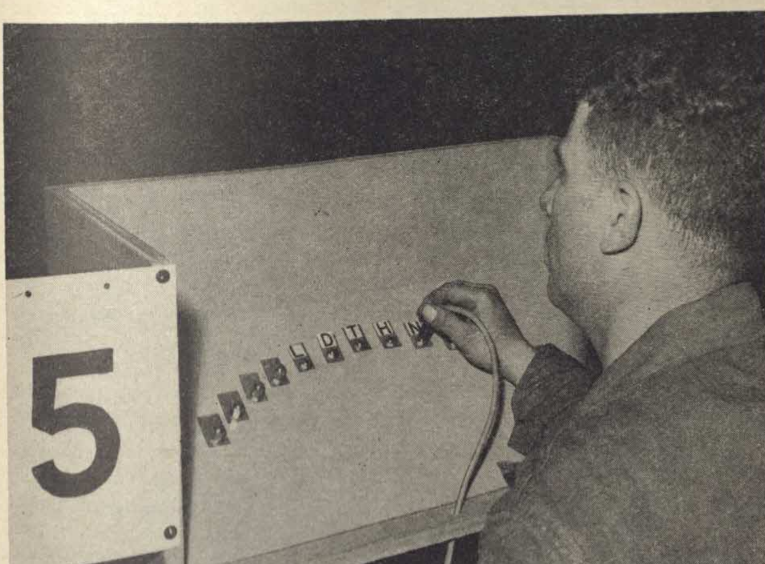
Method and Procedure

Subjects and apparatus. The Ss were 384 young men engaged in basic infantry training at the U.S. Army Training Center, Armor, Fort Knox, Kentucky. No S served in more than one condition. They were seated at individual response-panels in squads of 12. Each condition was administered to two squads of 12, hence the total N

[4] L. M. Schipper, Prediction of critical events in contexts of different numbers of alternative events, *J. exp. Psychol.*, 52, 1956, 377-380.



A



B

FIG. 1. EXPERIMENTAL ROOM AND APPARATUS

(A) Over-all view. During experimental sessions the only lighting in the room was furnished by two floor lamps directed toward the ceiling; one between Stations 1 and 7 and the other between Stations 6 and 12. The dim illumination, design of the response-cabinets, and the layout of the room effectively prevented the *Ss* from observing others' choices. (B) Close-up of an *S*'s response-cabinet.

serving in each condition was 24. The apparatus, shown in Fig. 1, A and B, is described in an earlier study.[5]

Procedure. The events consisted of block letters of the alphabet projected one at a time on the front wall of the experimental room. The *S*s were instructed to predict on every trial the next letter that would be projected on the screen. They were urged to make as many correct predictions as possible and also to make their predictions as rapidly as possible. On each response-panel appeared the letters that were available as categories of choice. Under each letter on the panels there was a standard ¼-in. telephone jack and each panel was provided with a single plug. A trial began with the darkening of the projection screen which was the signal for each *S* to make his next choice by inserting his plug into the jack under the letter chosen. A trial ended with the projection of the letter on the screen, which was the signal for all plugs to be withdrawn in preparation for the next trial. The single plug prevented *S* from making more than one choice per trial. A uniform pace of 4-sec. per trial was maintained for all but the earliest trials.

TABLE I

DESIGN OF THE EXPERIMENT

Block	Program variables			No. choices available					
	No. of categories	$\pi(A)$	Stimulus-letters	2	3	4	5	7	8
I	2	.70	L, S	X	X	X		X	
II	2	.60	T, H	X	X		X		X
III	3	.70	L, S, D		X	X	X	X	
IV	3	.60	T, H, N		X	X	X		X

Design. The 16 conditions of the experiment were divided into four Blocks of four conditions each according to the scheme shown in Table I. Within each Block the program of events that was presented was identical for all four conditions. For one condition in each Block the available choices matched the categories of events in the program, while for the remaining three conditions the available choices consisted of the categories in the program plus additional categories that were never presented, *i.e.* the dummy choices. In Block I and Block III, $\pi(A)$ was 0.70, and in Block II and Block IV, $\pi(A)$ was 0.60. Only two categories of events were presented in the programs of Block I and Block II, and dummy choices were added in such a way that the total number of choices for each condition would correspond to the values of this variable that were used in a previous experiment.[6] There were three categories in the programs of Block III and Block IV, which were run three months later, and dummy choices were added to parallel the previous work within the restrictions of the over-all design. In effect every Block was a self-contained experiment designed to provide a direct comparison between the dummy-choice procedures of this study and the multiple-choice procedures of the previous study. To accomplish this within certain practical limitations, it was necessary to sacrifice the possibility of direct cross-comparisons between the Blocks of the present study.

[5] Gardner, *op. cit.*, 1958, 710.
[6] Gardner, *op. cit.*, 1958, 711.

Spatial counterbalancing. The letters that represented the categories of choices are shown in Table II. This table also indicates the scheme by which the spatial arrangement of the letters on the response-panels was partially counterbalanced. Spatial arrangement of the letters appearing in the three center positions was completely counterbalanced. That is, each of the six possible arrangements was used for an equal number of *S*s. For conditions requiring Positions 4 to 9, the additional letters were arranged as shown in Table II for half of the *S*s while the mirror images of these arrangements were used for the other half.

Program. A table of random numbers was used to construct the programs. There were four programs, one for each Block, and the letters and values of $\pi(A)$ are shown in Table I. The most frequently presented letter in the programs for Block I and Block III was *L* for half of the *S*s in each condition and *S* for the other half. Similarly, *T* and *H* were alternated as the most frequent letter in the programs of

TABLE II

SPATIAL ARRANGEMENT OF RESPONSE-PANELS

Blocks I and III ($\pi(A)=0.70$)

Total choices	panel positions: left 6	left 4	center 2	center 1	center 3	right 5	right 7
2			L	S			
3 (I)			L	S	N		
3 (III)			L	S		D	
4			L	S	N	D	
5		H	L	S	N	D	
7	B	H	L	S	N	D	T

Blocks II and IV ($\pi(A)=0.60$)

Total choices	panel positions: left 8	left 6	left 4	center 2	center 1	center 3	right 5	right 7	right 9
2				T	H				
3 (II)				T	H	D			
3 (IV)				T	H		N		
4				T	H	D	N		
5			L	T	H	D	N		
8		B	L	T	H	D	N	S	F

Block II and Block IV. Each program was 140 trials long and the relative frequencies required by the design were fixed for these 140 trials although random fluctuations were permitted within sequences of less than 140 trials. For each condition, this program was run forward for the first 140 trials, reversed for the second 140 trials, and then run forward again for the third 140 trials, making a total of 420 trials.

RESULTS

Because the conditions stated in the design were only met in the programs for sequences of 140 trials, this was used as the unit for analysis of the $P(A)$ data. These data were tabulated in three periods of 140 trials; Period 1, Trials 1–140; Period 2, Trials 141–280; and, Period 3, Trials 281–420. The multiple-choice effects that have been demonstrated seem to depend upon the number of trials hence it is the data of the later periods that are critical for Hypothesis 1, and these are shown in Table III.[7] In all of the statistical tests quoted, the $P(A)$ score for each *S* was taken as the number of choices of *A*, the most frequently presented category, that

[7] Gardner, *opp. citt.*, 1957, 177-179; 1958, 714-715; Cotton and Rechtschaffen, *op. cit.*, 96.

were made during the Period in question. For convenience the values presented in Table III are the means of these scores converted into proportions of 140 trials. A separate 4 x 2 analysis of variance was computed for each Block for each Period using the number of choices as one variable and the letters representing the *A* category (*L* vs. *S* for Blocks I and III, *T* vs. *H* for Blocks II and IV) as the second variable. No *F* computed for *L* vs. *S* or *T* vs. *H* or their interaction with the number of choices was significant for either Period. The *F*s computed for the number of choices appear in Table III.

By Hypothesis 1, $P(A)$ should have increased with the number of dummy choices, at least by the last Period. By Period 3, of the 12 independent conditions with dummy choices, only one failed to reach a $P(A)$ higher than that of the condition with the same program but no dummy choices. Within each Block the rank order of $P(A)$ during Period 3 agrees with the rank order of $P(A)$ with only two exceptions, and the highest $P(A)$ within each Block was that of the condition with the largest number of dummy choices. The strength of this trend was statistically significant, however, only in the case of Block III. When differences between pairs of conditions within Blocks were tested, only *t*-tests for comparisons in Block III were significant.

TABLE III

$P(A)$ FOR PERIOD 2 (TRIALS 141–280) AND PERIOD 3 (TRIALS 281–420)

Block	Period	Total choices				F
		2	3	4	7	
I	2	.709	.754‡	.767‡	.752‡	1.11
	3	.730	.756‡	.759	.763‡	0.27
		2	3	5	8	
II	2	.584	.559	.587	.573	0.36
	3	.596	.562	.609	.628†	1.19
		3	4	5	7	
III	2	.713	.710	.772‡	.809‡	2.81*
	3	.725	.737†	.811†‡	.834†‡	3.81*
		3	4	5	8	
IV	2	.585	.609	.656	.619	1.15
	3	.623†	.680†‡	.676‡	.687†‡	0.89

* $P<0.05$.
† Period 3 > Period 2 by paired-differences *t*-test with $P<0.05$.
‡ $P(A)>\pi(A)$ by *t*-test using raw scores with $P<0.05$.

When the responses to the dummy choices were tabulated, it became more clear why they had such a weak effect. Less than 50% of the *S*s showed any response to the dummy choices after Trial 16, less than 12% after the beginning of Period 2 (Trial 141), and less than 7% after the beginning of Period 3 (Trial 281), A measure of the rate of extinction

was obtained by tabulating the trial on which each *S* made his last dummy choice, and computing the median for each condition and each Block of conditions. These medians are shown in Table IV. Within each Block, the greater the number of dummy choices, the earlier they were extinguished. Median tests were performed on these data for each Block and the resulting χ^2s are also shown in Table IV.[8] The χ^2s for Block I and Block II were significant, while those for Block III and Block IV were not. As there were no reversals in the direction of the effect, the data for all

TABLE IV

EXTINCTION OF THE DUMMY CHOICES

	No. dummy choices	1	2	3	Block
Block I	Median ext. trial	13.5	10.5	6.5	10.5
	*S*s above Block median	18	12	6	$\chi^2=12.00^*$
	No. dummy-choices	1	3	6	Block
Block II	Median ext. trial	48.0	12.0	10.0	14.0
	*S*s above Block median	17	9	8	$\chi^2=8.14^*$
	No. dummy-choices	1	2	4	Block
Block III	Median ext. trial	58.0	36.0	18.0	32.5
	*S*s above Block median	15	13	8	$\chi^2=4.33$
	No. dummy choices	1	2	5	Block
Block IV	Median ext. trial	30.5	26.0	14.0	23.5
	*S*s above Block median	15	12	9	$\chi^2=3.00$
				Totaled	$\chi^2=27.47\dagger$

* $P<0.02$, with 2 df.
† $P<0.01$, with 8 df.

four Blocks can be combined by adding the χ^2s.[9] The totaled χ^2 which is shown in Table IV is highly significant (27.47 with 8 df.).

Regarding Hypothesis 2a, by which $P(A)$ should be expected to level-off, it can be seen in Table III that $P(A)$ was higher during Period 3 than during Period 2 for 15 of the 16 independent conditions. This was evaluated by paired difference *t*-tests computed separately for each condition and the results in Table III show that these differences were significant for 7 of the 15 conditions. It might be argued that the same result might be obtained if $P(A)$ leveled-off anywhere during Period 2, that is, within 280 trials. That such leveling was not typical of these data is shown in Fig. 2, where $P(A)$ is plotted for each Block for the last two Periods, Trials 141–420. Regarding Hypothesis 2b, by which $P(A)$ would not be expected to exceed $\pi(A)$, it can be seen in Table III that during Period 3, $P(A)$ did exceed $\pi(A)$ for 14 of the 16 conditions. By *t*-tests, seven of these differences were significant as shown in Table III. It should be noted that the probability levels quoted in Table III refer to two-tailed tests.

[8] Sidney Siegel, *Non-Parametric Statistics for the Behavioral Sciences*, 1956, 179-184.
[9] G. W. Snedecor, *Statistical Methods*, 1956, 212-236.

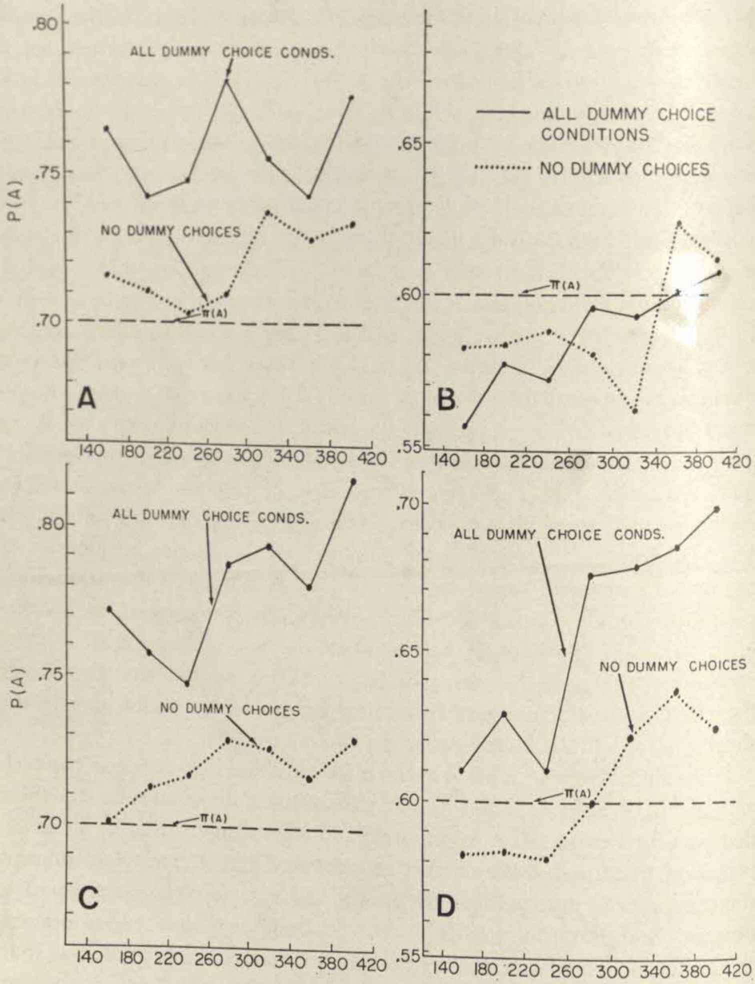


FIG. 2. $P(A)$ PLOTTED FOR EVERY BLOCK DURING THE LAST TWO PERIODS
(Every point is the average of 40 trials.)
(A) Block I; (B) Block II; (C) Block III; (D) Block IV

DISCUSSION

In the previous experiments it was possible to obtain clearly significant enhancement of $P(A)$ as a function of the number of choices when the non-A events in the program were equally divided among the added choices.[10] Hypothesis 1 was framed as an interpretation of this enhancement as a function of the number of available choices rather than the num-

ber or relative frequencies of the categories in the program. In the present experiment, adding categories of choices without adding categories of events in the program tended to raise $P(A)$. Such effects as appeared, however, were clearly not comparable to those obtained under the more conventional multiple-choice procedures with the possible exception of the conditions of Block III. At most, these results provide only very weak support for Hypothesis 1; but, the dummy choices were so readily extinguished that, perhaps, only weak effects are to be expected with this technique. While the dummy choices attracted so little response that it is difficult to evaluate Hypothesis 1, the rate of extinction proved to be sensitive to the number of dummy choices, and the direction of this effect suggests a very meaningful alternative. Within each Block of conditions, the more dummy choices available, the more rapidly all response to dummy choices was extinguished. This is precisely the result that would be expected if the multiple-choice effect is seen as a depression of response to the less frequent categories rather than an enhancement of response to the most frequent category. Previously, this distinction was not clear because only $\pi(A)$ was held constant, while both the number and the relative frequencies of the non-A categories varied together. By the technique of the present experiment, for all conditions within a Block, the program of events was identical and, regardless of their number, the rate of hits for all dummy choices was the same, *i.e.* zero. The rate of misses was the same also, hence the effect emerges clearly as an accelerated rate of extinction due to the sheer number of dummy choices on the response-panels.

It should be kept in mind that the method of scoring extinction for each S might have been expected to bias the scores, if at all, against the effect that was demonstrated. Also, the trends in $P(A)$ might seem to fit an explanation in terms of the number of choices diverted from A to dummy choices. In most groups, however, no choices at all were being diverted to dummy choices during Period 3, and in those few cases where dummy choices were being made during Period 3, the total number was less than 1% of all the responses of the group. Furthermore, if $P(A)$ were lowered by the diversion of choices to the dummy choices, then the condition within each Block which had no dummy choices available should have had the highest $P(A)$ in each Block instead of the lowest.

Hypothesis 2 a is concerned with the question of whether S will eventually respond in such a way as to maximize the number of hits he obtains so that $P(A) = 1.0$, or whether his responses are so governed by factors other than the maximization of hits that $P(A)$ will level-off at some as-

[10] Gardner, *opp. citt.*, 1957, 177-179; 1958, 714-715; Cotton and Rechtschaffen, *op. cit.*, 96.

ymptote below 1.0. The first alternative is not clearly confirmed unless a $P(A)$ of 1.0 is actually obtained. Nevertheless, it cannot be clearly rejected in favor of the second alternative merely because, in a given experiment, $P(A)$ fails to reach 1.0 after what seems to be a reasonable number of trials. The confidence with which the value of $P(A)$ during the terminal trials of a given experiment can be accepted as an estimate of some theoretical asymptote depends entirely upon the stringency of the criterion used to rule out the possibility of further increments. While no really conclusive statement can be made here regarding asymptotic response within 420 trials, when the results of the present experiment are taken together with those of a similar experiment performed earlier,[11] it seems evident that asymptotes were not reached within 280 trials, with the possible exception of two-choice conditions. On the other hand, Hypothesis 2 b, the matching hypothesis, can be rejected with some confidence. With multiple-choice conditions values of $P(A)$ that significantly exceed $\pi(A)$ can readily be obtained within 420 trials.[12] Edwards has shown that a patient experimenter can obtain values of $P(A)$ that exceed matching, even with two choices, by extending the experiment for a sufficient number of trials.[13]

SUMMARY

A Humphreys-type verbal conditioning-procedure was used to study multiple-choice behavior in an uncertain situation. The number of choices was varied by adding dummy choices to the categories of choice that were available to *S*. These dummy choices represented categories of critical events that never appeared in the programs. Under these condtions, there was a tendency for *S*s to choose the most likely alternative more often as a function of the total number of available choices, but this trend was not statistically reliable. Response to the dummy choices extinguished so rapidly, however, that an unequivocal interpretation of this result was not possible. The rate of extinction was significantly related to the number of dummy choices. The more dummy choices there were available the more rapidly all response to the dummy choices was extinguished. This result suggests that multiple-choice behavior in uncertain situations might be understood in terms of the extinction of competing alternatives. To the extent that the data were relevant to the hypothesis of asymptotic matching, it was not confirmed.

---

[11] Gardner, *op. cit.*, 1958, 714-715.

[12] P. D. McCormack, Spatial generalization and probability-learning in a five-choice situation, this JOURNAL, 72, 1959, 135-138; Gardner, *opp. citt.*, 1957, 177-179; 1958, 714-715; Cotton and Rechtschaffen, *op. cit.*, 96.

[13] Ward Edwards, Probability learning in 1000 trials, *Amer. Psychologist*, 13, 1958, 424 (Abstract).

RECENCY AS A FUNCTION OF PERCEPTUAL OSCILLATION

By WILLIAM EPSTEIN and DEL DE SHAZO, University of Kansas

In a recent paper Epstein and Rock reported the results of an experiment on the influence of expectancy on the perception of form.[1] Expectancy was separated from those factors which are usually coincident with it, and the relative effectiveness of these factors was tested in situations where they would lead to mutually antagonistic perceptual outcomes.

Thus, in one experiment *S* was first shown the two unambiguous profiles of the ambiguous Schafer-Murphy figure-ground composite.[2] *S* was then informed that *E* would present a series of four slides consisting of three representations of *A* and a single *B*. After three consecutive tachistoscopic presentations of *A*, almost all *S*s expected to see *B*. Instead, *A/B* was presented tachistoscopically. Despite an expectancy for *B*, 71% of the *S*s perceived *A* which was consonant with the most recent perceptual experience.[3] Similar results were obtained with a modified version of the ambiguous figure, 'My wife and my mother-in-law'.[4]

These experiments led Epstein and Rock to conclude that the primary determinant of perception in many preparatory situations is a specific, recently formed memory trace and not expectancy as such. This conclusion confronts us with the logical problem of trace-selection which is present whenever we consider an effect of memory on perception.[5] Appropriate selection of traces seems to depend on some kind of similarity between the present perceptual process and the memory trace deposited by the previous process. This means that the present perceptual process must be organized

* Received for publication January 20, 1960.

[1] William Epstein and Irvin Rock, Perceptual set as an artifact of recency, this JOURNAL, 73, 1960, 214-228.

[2] Roy Schafer and Gardner Murphy, The role of autism in a visual figure-ground relationship, *J. exp. Psychol.*, 32, 1943, 336 ff. To facilitate discussion, we shall refer to the ambiguous composite as *A/B*. The profile which in the usual reproduction occupies the left half of the composite facing the right will be called *A*. The profile occupying the right half facing left will be called *B*. See Fig. 1.

[3] In this experiment frequency may also be a factor in determining the outcome. This variable was investigated in other experiments.

[4] E. G. Boring, A new ambiguous figure, this JOURNAL, 42, 1930, 444-445.

[5] For a more complete discussion of this problem see Hans Wallach, Some considerations concerning the relation between perception and cognition, *J. Person.*, 18, 1949, 6-13; C. B. Zuckerman and Irvin Rock, A reappraisal of the roles of past experience and innate organizing processes in visual perception, *Psychol. Bull.*, 54, 1957, 269-296, esp. 277-279.

before it can communicate with the trace since only an organized process (*i.e.* one resulting in a definite shape in the case of form perception) can be similar to the trace representing the organized form seen previously. It appears, therefore, that the trace cannot enter into the perceptual process until the process itself is organized. Once this has occurred, however, it is no longer necessary to invoke a trace effect to explain the perceptual outcome.

Since the effect of recency is obviously an influence of past experience, we are faced with the paradox of trace-slection. The difficulty is not so acute in regard to the effect obtained with the Boring figures. It is very likely that in this case the selection of the trace was mediated by some partial similarity between an aspect of the composite in an early stage of development and the trace of the recent, unambiguous percept. On this

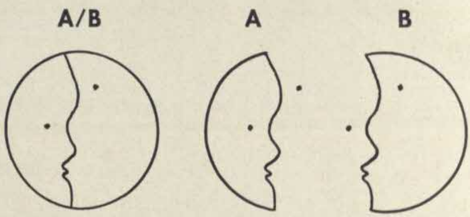


FIG. 1. THE SCHAFER-MURPHY AMBIGUOUS COMPOSITE AND THE TWO SEPARATE UNAMBIGUOUS PROFILES

basis the trace could enter into the developing process and influence the organizational outcome. Such considerations, however, will not apply to the effects obtained with the Schafer-Murphy figure. In this instance the only point of similarity which can serve to mediate trace selection is perceived shape itself.

The resolution which we propose is essentially an effort to circumvent the problem. We contend that the influence of recency does not reside in the determination of the organization of form *per se,* but that the recent trace determines the selection and stabilization of perceptual processes *already organized.*

The major premise underlying this proposal is the assumption of perceptual oscillation. We assume that when the labile, ambiguous stimulus-configuration (*A/B*) is presented under marginal perceiving conditions, there occurs a rapid oscillation of the perceptual alternatives prior to the exclusive emergence of one of the alternatives. *A* and *B* occur as organized forms in rapid alternation below the threshold of awareness. When the

alternative appears which is congruent with the recently implanted trace, then arousal of the trace occurs on the basis of distinctive similarity. The trace, once selected, intrudes into the labile perceptual process and stabilizes the situation by reinforcing the consonant perceptual alternative. As a consequence of these presumed occurrences, recency exerts a selective influence, determining which of the already organized alternatives will gain dominance and be perceived. In this way the logical issue is avoided since it is no longer necessary to assume the selection of traces prior to the organization of the consonant percept. Instead, the trace is contacted *after* the percept is organized, and only then does it exert a selective influence.

Method. There are many difficulties attending an experimental test of this hypothesis. We cannot look to *S*'s introspective report for information about the occurrence of oscillation. Furthermore, the oscillation is assumed to occur so rapidly as to preclude an attempt to 'break into' the ongoing process.

We finally decided upon the following procedure. A discriminatory *GSR* would be conditioned to one of the alternatives under circumstances that favor a recency-effect of the other alternative. For example, *A* would serve as the *CS* while *B* would be most recent, *i.e.* immediately precede *A/B*. When *A/B* is presented, *S* should see *B*, but, at the same time, the *GSR* which is conditioned to *A* only should be evoked. This follows from our assumption that prior to the exclusive appearance of *B*, perceptual oscillation took place in which *A* was also represented.

Apparatus. A La Belle 2 × 2-in. slide projector was used. An Alphax tachistoscopic lens calibrated for time-exposures ranging from 1/100 to 1 sec. was attached. A 50-slide cartridge and semi-automatic feeding mechanism permitted exposures to be made quite rapidly. The images were projected onto a white screen 65 cm. from the projector and approximately 45 cm. from *S*. The room was dimly lighted. A shock device, constructed from a Variac transformer (range of 0 to 130 v.), was attached to the ring and index fingers of *S*'s left hand and was so designed that a current of variable intensity could be applied by pressing a 'silent' switch. A Hunter model 300-*GSR* amplifier was connected to both the palm and back of *S*'s right hand by two electrodes, well saturated with electrode jelly. The *GSR*-recorder was connected in series with an Esterline-Angus Graphic Ammeter (model AW) and was calibrated so that each ¼-in. of arc swing by the pen recorder represented the addition (positive swing) or reduction (negative swing) of 250 Ω of resistance.

Materials. The Schafer-Murphy composite, *A/B*, and the two unambiguous profiles, *A* and *B*, were used. The fourth figure was a circle, *C*, identical in circumference to the border of the composite. These figures were drawn separately on 2 × 2-in. transparencies.

Subjects. The *Ss* were 71 students (35 men and 36 women) in an introductory course in psychology. All of them were naïve regarding the purpose of this study and the nature of the perceptual stimuli.

Procedure. The *Ss* were tested individually in a soundproof room. Each in turn was seated at a table containing the apparatus previously described, and the purpose of each piece of equipment was explained briefly. The shock-apparatus was attached to *S*'s left hand and shock was administered beginning with an intensity of 5 v. and then increased by increments of 5 v. This continued until *S* found the stimulus to be disagreeable. When this level was established the electrodes of the amplifier were attached to *S*'s right hand.

Instructions. *S* was instructed to relax for several minutes during which time his 'basic resistance-level' was recorded. Then the following instructions were read to him.

In this experiment I am trying to discover how rapidly different people can identify different visual figures, even when they are occasionally distracted by electric shock. Three different figures are going to be flashed on the screen many times in random order for very brief exposures. [At this point Stimulus-Figure *A* was described and presented briefly; Stimulus-Figures *B* and *C* were then described in turn and presented briefly. The instructions were then continued as follows:] Do you think you can recognize these figures? [If the answer was negative, each figure was described verbally again.] Now we are ready to begin the experiment. As soon as a figure appears on the screen you are to identify it by calling out *immediately* whether it is *A*, *B*, or *C*. Remember, call out the letter assigned to the figure the *instant* you see it! Ready? Let's begin.

A preparatory series of 46 slides was presented. The following conditions were imposed on the composition of the series. (1) *A* and *B* occurred with equal frequency, 19 presentations each. (2) *A* and *B* faced both left and right an equal number of times. (3) *A* and *B* never followed each other in complementary orientations on successive presentations, *e.g.* *A* facing right followed by *B* facing left. (4) The final *A* or *B* (Presentation 46) which immediately preceded *A/B* faced in a direction opposite to the direction of that profile in the composite. (5) *C* was presented eight times. In other respects the preparatory series was ordered randomly. This series was followed by the presentation of *A/B*.

Slides 1–42 each were presented for 0.02-sec.; Slides 43–46 and *A/B* each were presented for 0.01 sec. The shorter exposure-time for *A/B* was intended to reduce the number of 'both' responses, and it was initiated with Presentation 43 so as to allow sufficient time for the extinction of any *GSR* which might result from the sudden introduction of a more rapid exposure. All exposures, except for every tenth, followed the preceding one by 5 sec.; each tenth exposure followed the preceding one by only 1 sec. In this way we reduced the possibility that the 1-sec. interval between the composite and the immediately preceding figure would elicit a *GSR* in reaction to the shorter interval.[6]

All incorrect identifications of Presentations 1–7 (3 *A*s, 3 *B*s and 1 *C*) were cor-

---

[6] *C* was introduced into the preparatory series for the same reason. We tried in this manner to eliminate a *GSR* to *A/B* which might be produced by its novel circular appearance. Almost without exception these precautions proved to be unnecessary. A significant change in *GSR* rarely was observed to accompany the initial or subsequent occurrences of these novel stimulus-conditions.

rected by *E*. For Group 1, 15 of the remaining exposures of *B*, the conditioned stimulus (*CS*), were followed after 0.5 sec. by a shock, the unconditioned stimulus (*US*). The final presentation of *B* was not followed by shock. This was a test-trial. The occurrence of a *GSR* on this trial was taken as evidence that a conditioned response (*CR*) was established. Group 2 was treated similarly except that *A* served as the *CS*. For Group 1, *A* was most recent, *i.e.* preceded *A/B* and for Group 2, *B* was most recent.

The following criterion was adopted to determine whether an obtained recency-response was accompanied by a conditioned *GSR*. The *GSR* elicited during the test-trial for a given *S* served as an indicator of the upper limit of *GSR* magnitude for the next appearance of the non-recent figure. A *GSR* of somewhat lesser intensity would be expected to accompany a recency-response to *A/B* in cases where the non-recent profile had appeared in the prior oscillation. The *GSR* elicited by the final presentation of the recency-figure was taken to represent the lower limits of any *GSR* to *A/B*. This is the response-level to be expected if only the recency-figure is perceived when *A/B* is presented. To qualify as a *CR* any *GSR* accompanying a recency-response to *A/B* had to exceed the mean of the response-magnitudes for the final presentations of the recent and non-recent figures.

Results. Eleven *S*s were rejected from the study because they failed to develop a *CR*. All the remaining *Ss* developed the required discriminatory response, *i.e.* a significant change in the *GSR* was elicited by the *CS* and was not elicited by either of the two neutral stimuli. The results which are reported below are based on two groups of 30 *S*s each. Since there was no difference in the responses for the two groups, we have combined the data for both groups, *i.e.* for 60 *S*s.

(*1*) The first aspect of our results with which we are concerned is the frequency of the recency-responses, *i.e.* the number of *S*s who saw *A/B* in accordance with the most recent perceptual experience. The responses were distributed as follows: 36 *S*s saw the recency-figure, 9 saw the non-recent alternative, and 15 saw both. If we exclude 'both' responses, we find that 80% of the unambiguous responses were consonant with recency ($X^2 = 16.2$, $p < 0.01$). These results confirm the earlier findings reported by Epstein and Rock.[7]

(*2*) The second aspect of concern is the frequency at which the conditioned *GSR* accompanied the recency-responses made to *A/B*.

Our hypothesis does not demand that all the recency-responses be accompanied by the conditioned *GSR*. In fact, as a first approximation we might predict that only 50% of the recency-responses will be accompanied by a *CR*. This expectation follows from the assumption that during preliminary oscillation the alternative which is consonant with recency, *i.e.* the

[7] Epstein and Rock, *op. cit.*, 214-228.

neutral stimulus, will appear first in half the cases. When this happens the oscillation will be aborted. Trace-arousal will occur and the consonant percept will be immediately stabilized, thereby precluding the emergence of the *CS*, *i.e.* the alternate profile, and the associated *CR*.

On the basis of the assumption given immediately above we may sharpen our prediction. If the 'both' responses (15) are eliminated, and the remainder (45) is halved, then the quotient may be taken to represent the number of occasions on which the non-recency figure appeared first in the oscillation process, *i.e.* 50% of the cases or 22.5. From this assumed value we may then subtract the obtained number of non-recency responses (9). The remainder may be presumed to represent the number of instances when the initially occurring non-recency percept was not reported because it was superceded by the influence of recency, *i.e.* the percept favored by recency was attained. On these occasions the recency-response should be accompanied by a conditioned *GSR*. This prediction is most easily expressed by the following simple formula:

$$GSR_{A/B} = [(T - B)/2] - NR$$

where the left hand term is the frequency of *GSR* occurrence when A/B is presented, T is the total number of cases, B is the number of 'both' responses, and NR is the number of non-recent responses. We may substitute our present results in the formula as follows:

$$GSR_{A/B} = [(60 - 15)/2] - 9.$$

The value 13.5 which is obtained in this way is the theoretically expected frequency of *GSR* occurrence. This value represents 37.5% of the total number of recency-responses. We may summarize this reasoning by saying that in 37.5% of the cases where a recency-response is given, a *GSR* will be evoked also.

Our results fit this requirement. The data show that 14 (38.9%) of the 36 recency-responses were accompanied by the conditioned *GSR*. This frequency does not differ significantly from the expected frequency of 13.5 (37.5%).[8]

Discussion. We have presented evidence in support of our assumption that subthreshold, perceptual oscillation does occur. In the preceding experiment the conditioned *GSR* was elicited when A/B was presented with a frequency which satisfied the theoretical requirement. This is a fairly

[8] As would be expected, all of the non-recency responses and 'both' responses made to A/B were accompanied by *GSR*.

unequivocal indication that the *CS*, *i.e.* the non-recent alternative, did occur prior to the final emergence of the alternative favored by recency.

There are, however, other possible interpretations of our data which should be mentioned:

(1) Stimulus-generalization. The *GSR* elicited might be the result of stimulus-generalization from the *CS* to the neutral stimulus (recency-figure). This possibility cannot be taken too seriously since the preparatory series was very successful in establishing a discriminatory response to the *CS*. A careful inspection of the graphic records of the *GSR* convinced us that any original tendency toward generalization was completely extinguished long before the test-presentation.

(2) Response inhibition. The occurrence of a *CR* without the accompanying report of the *CS* might be attributed to a suppression of the verbal report. In this view, *S* defends himself against the noxious potential of the *CS* by suppressing any *overt* recognition of its existence. Thus a *CR* is obtained, but a verbal report is omitted. There are many reasons for discounting this interpretation. Most importantly, it can be dismissed because it cannot account for our results. On the basis of this thesis no prediction can be made about the proportion of recency-responses which will be accompanied by a *GSR*.[9] Nor is this hypothesis compatible with the distribution of test-responses which was obtained. One would expect that as part of a defensive effort every *S* would give a recency-response. This was not the case.

(3) Subjective probability of CS. Another explanation of our findings might be derived from Howes' analysis of the effect of subception.[10] Howes interprets the results obtained in the experiments on subception in terms of the hypothesis that "at any specified moment the *GSR* accompanying an observer's report is proportional to the probability that that report will be a shock (conditioned) syllable."[11] Applying this hypothesis to the present experiment, one might argue that *S* expects shock to accompany a third of the presentations. Therefore, when *A/B* is presented a third of the *Ss* expect shock, and consequently a *GSR* is elicited in a third of the cases.[12]

This hypothesis is not consistent with the performances during the training series. If the occurrence of a *GSR* is correlated with the subjective

[9] The same objection applies to the first alternative discussed above.
[10] Davis Howes, A statistical theory of subception, *Psychol. Rev.*, 61, 1954, 98-110.
[11] *Ibid.*, 99.
[12] In point of fact our interviews rarely revealed that the *S* had any specific expectancy about the identity of the coming event.

probability that the approaching presentation will be the *CS*, then there should have been many instances when a *GSR* accompanied the two neutral stimuli. This, however, seldom happened. After the twenty-fifth presentation, none of the presentations of the two neutral stimuli was accompanied by a significant change in *GSR*.

The primary purpose of this investigation was to provide a more complete understanding of the processes underlying the effect of recency. At the same time, we have demonstrated one of the ways in which a trace can enter into an ongoing perceptual process and influence its further development. No doubt the thesis of perceptual oscillation contributes only a partial resolution of the paradox of trace-selection and will not apply to all illustrations of the effects of past experience in perception. We have indicated earlier that the effects obtained by Epstein and Rock with the "Wife and Mother-in-Law" figure are probably to be explained differently.[13] The experiments performed by Wallach, O'Connell, and Neisser in which a memory-trace is shown to affect depth-perception in a *KDE* (kinetic depth effect) situation also may be attributed to different processes.[14]

Our formulation, however, may have general applicability where relatively simple, labile stimulus-configurations are employed. Thus, in a recent study, an effect of past experience was demonstrated on the perception of a reversible, tridimensional cube.[15] *S* inspected a skeletal wire cube monocularly until a reversal occured, *i.e.* an inversion of the perspective of the near and far faces. Both eyes were then covered for a 2-sec. period after which *S* looked monocularly at the cube again. The influence of past experience is shown by the significant number of cases in which the reversed version of the cube was seen after training. The applicability of our formulation to this situation is fairly apparent.

Another significant aspect of the present study is the consistency with which the recency-effect occurs. For five experiments (including the experi-

---

[13] *Op. cit.,* 214-228; Also of interest are the results reported by Robert Leeper (A study of a neglected portion of the field of learning—the development of sensory organization, *J. genet. Psychol.,* 46, 1935, 41-75). Leeper, using the "Wife and Mother-in-Law" figure, was able to show an effect of previous perceptual experience after an interval of two weeks. This is further indication that the effects obtained with that figure and the Schafer-Murphy figures are not to be explained in the same way.

[14] Hans Wallach, D. N. O'Connell and Ulric Neisser, The memory effect on visual perception of three-dimensional form, *J. exp. Psychol.,* 45, 1953, 360-368. There are, however, certain interesting similarities between the explanation which Wallach *et al.* offer and our own.

[15] P. A. Adams, The effect of past experience on the perspective reversal of a tridimensional figure, this JOURNAL, 67, 1954, 708-710.

ments reported by Epstein and Rock) using over 400 *Ss* from three different populations, the magnitude of the recency-effect has ranged from 71%–80% of the unambiguous responses. The demonstrated reliability of this effect may perhaps make it a useful tool for the indirect investigation of the nature of the memory-trace.

Finally, we wish to mention briefly one tangenitial aspect of our study. We believe that the results of this experiment constitute a demonstration of subception which is not open to the methodological criticisms often directed at previous experimental illustrations of this phenomenon.[16] In addition, our experiment has the advantage of being free of any confounding association with the concept of perceptual defense.

SUMMARY

In a previous paper it was shown that many of the perceptual effects usually attributed to expectancy should be attributed instead to the influence of specific, recent, memory-traces. This conclusion made it necessary for us to explain how a memory-trace can enter into a perceptual process and influence its outcome. The hypothesis which we proposed centered around the assumption that prior to the exclusive emergence of the percept finally attained, there occurred a period of perceptual oscillation. During oscillation the various organized alternatives appear and trace-selection occurs on the basis of distinctive similarity. Once the trace is selected, it can then enter into the labile process and determine its further development.

This hypothesis was tested in an experiment utilizing a classical conditioning procedure, and supportive evidence was obtained. The significance of this finding was discussed with reference to several questions in the area of perception.

[16] See especially C. W. Eriksen, Subception: Fact or artifact, *Psychol. Rev.*, 63, 1956, 74-80; Discrimination and learning without awareness: A methodological survey and evaluation, *ibid.*, 67, 1960, 279-300.

THE EFFECT OF SHIFT OF SENSORY MODALITY ON SERIAL REACTION-TIME: A COMPARISON OF SCHIZOPHRENICS AND NORMALS

By SAMUEL SUTTON, GAD HAKEREM, and JOSEPH ZUBIN, Columbia University, and MAURICE PORTNOY, Brooklyn State Hospital

Information about the environment is available to the organism by way of a number of simultaneously active sensory channels, but all of these inputs do not have equal priority at any given moment. There has, however, been relatively little study of the dominance of any one specific sense and the consequences of this dominance for behavior. Nor do we have experimental information regarding the ease with which this relative dominance may be shifted from one sense to another. Interest in this problem, however, dates back to Wundt, who commented as follows.

> Slighter but still very noticeable is the retardation [in quickness of response] if one arranges the experiment to have the observer in ignorance as to whether light, sound, or touch impressions [stimuli] will be forthcoming, so that the attention cannot be turned to a particular sense organ. Immediately one notes a peculiar unrest because the strain of attention [*die Aufmerksamkeit begleitende Spannungsgefühl*] continuously vacillates among the several senses.[1]

More recently Mettler pointed to the role of this "vacillation of attention" among the several senses in environmental scanning by the organism.[2] He suggests that the relative facility of shifting set or attention from one sensory input to another may be relevant to the efficiency of the organism's processes of monitoring events in progress. He speculates that deficiencies in maintenance of set, and deficiencies in capacity to shift set, may account for some aspects of the perceptual disorientation of the schizophrenic. He goes on to suggest specific brain structures whose malfunction might cause impairment in the control of information coming in along competing sensory routes.

Mettler's neural hypotheses need not be fully accepted at this time but they suggest, nevertheless, a comparison of schizophrenic patients and

* Received for publication October 30, 1959. This investigation was supported in part by Projects M 586 and M 1541 from the National Institute of Mental Health granted to Biometrics Research, New York State Department of Mental Health.

[1] Wilhelm Wundt, *Grundzüge der physiologischen Psychologie,* 4th ed., 2, 1893, 352 f. (Translation ours.)

[2] F. A. Mettler, Perceptual capacity, functions of the corpus striatum and schizophrenia, *Psychiat. Quart.,* 29, 1955, 89-109.

normal individuals regarding the relative ease with which they may shift a response to stimuli in one sensory modality to stimuli in another modality. We expected schizophrenics to have more difficulty than normals in this shift.

For the purpose of this study, we used the simple task of lifting the finger from a 'home' plate as rapidly as possible after the onset of a light or a sound. We made the assumption that having just responded to one stimulus, *S* is set to respond to another stimulus in the same modality more readily than to a stimulus in another modality. As a measure of the ease of shifting, we used the relative speed of response to a second stimulus in the same modality as compared with a stimulus in a different modality. We shall call the former the 'ipsimodal' reaction-time and the latter the 'crossmodal' reaction-time.

Method and Procedure

Many experiments which attempt to compare the level of performance of schizophrenics and normals suffer from the difficulty that differences between groups not intrinsic to the purpose of the experiment probably account for most of the observed differences in performance. For example, it has been widely observed that schizophrenics display longer reaction-times than normals, but it is not clear whether this represents some kind of motor blocking, or simply reflects lack of coöperation and poor motivation on the part of the schizophrenic patient.[3] In the design of the present experiment, we attempted to minimize the effect of these extraneous variables by making each *S* his own control. A preliminary description of our experimental procedure may serve to clarify our approach to this problem.

S was instructed to lift his finger as rapidly as possible at the onset of any one of the four stimuli (red light, green light, high tone of 1200~, low tone of 400~) used in the series. When he had accomplished this, he returned his finger to the home plate and waited for the next stimulus. The stimuli were programmed in a quasi-random order and were separated in time by a few seconds. In the analysis of the data, we compared crossmodal reaction-time (*e.g.* reaction-time to a high tone which had been preceded by a trial with a red or green light) with ipsimodal reaction-time (*e.g.* reaction-time to a high tone which had been preceded by a trial with a low tone).

A lapse of motivation for part of the series or even for the whole series of stimuli should have affected both types of trials equally since they were interspersed in the presentation. By interweaving, moreover, the two experimental conditions, no special attention was called to the comparisons in which *E* was interested. The sequence of stimuli appeared purely random to *S*. Since the task is a simple reaction, requiring no discrimination in the perception of the signal nor in the execution of the response, the ipsimodal and crossmodal reaction-times should not differ unless the shift in modality *per se*—the phenomenon *E* was investigating—introduces a difference.

[3] H. E. King, *Psychomotor Aspects of Mental Disease*, 1954.

Subjects. The *Ss* for the experment were 23 chronic patients and 25 normal controls. The chronic patients had all been diagnosed as schizophrenic and had been residents of Broolyn State Hospital for at least two years.[4] Patients who had received tranquilizing drugs or other somatic treatment during the two weeks prior to the experimental session were excluded. The normal controls were obtained from personnel working in the hospital ward.

Procedure. Every *S* was tested on two successive days. The four stimuli originated at the same point in space in front of *S*. They were clearly discriminable and had been subjectively equated for intensity by one of the *E*s utilizing himself as an *S*. *S* was seated in front of the apparatus and made familiar with the four stimuli. He was instructed to lift his forefinger from a plate as rapidly as possible whenever any

TABLE I

SAMPLE FROM THE DATA FOR ONE *S* ILLUSTRATING THE MODE OF CLASSIFICATION USED IN THE STUDY

Trial No.	Stimulus	Code	Reaction-time (in m.sec.) Ipsimodal (identical)		Ipsimodal (non-identical)		Crossmodal	
			Light	Sound	Light	Sound	Light	Sound
1*	red	—						
2	red	rR	334					
3	red	rR	221					
4	low	rL						330
5	high	lH				365		
6	red	hR					224	
7	green	rG			187			
8	high	gH						175
9	high	hH		221				
10	red	hR					230	
11	high	rH						271
12	low	hL				198		
13	low	lL		184				
81								

* First trial not used.

of the four stimuli was presented. The lifting of the finger terminated the stimulus and *S* returned his finger to the home plate and waited for the next stimulus. Reaction-time was measured from the onset of the stimulus to the lifting of the finger. No ready-signal was given; the termination of each stimulus served as a signal that the next trial had begun. The time-intervals between trials were random values between 3–5 sec. automatically controlled by a prepunched film-strip moved by a constant-speed motor.

A series consisted of 80 trials, 20 for each of the 4 stimuli. The stimuli were presented in a mixed order so designed that there were as many crossmodal as ipsimodal sequences. No more than three stimuli in a given modality occurred successively. A part of the series is presented as an example to clarify the procedure and the classification used for analyzing the data.

As may be seen from Table 1, each trial may be classified as crossmodal or ipsi-

[4] Patients for whom the case record indicated any doubt with respect to diagnosis were eliminated from the sample.

modal, and as sound or light. The ipsimodal trials are further subdivided as having a stimulus identical or non-identical with the preceding stimulus. Since all crossmodal trials must be non-identical, there are only two groups of crossmodal trials, those to sound and those to light. For example, the stimuli of Trials 4, 8, and 11 (rL, gH, rH) are sounds which were preceded in each case by light. These trials are, therefore, crossmodal trials to sound. In like manner, Trials 6 and 10 (hR, hR) are crossmodal trials to light. Trials 2 and 3 (rR, rR) are ipsimodal identical trials to light, and Trials 9 and 13 (hH, lL) are ipsimodal identical trials to sound. Trial 7 (rG) is an ipsimodal non-identical trial to light, and Trials 5 and 12 (lH, hL) are ipsimodal non-identical trials to sound.

Analysis of data. For every *S*, we performed a Mann-Whitney *U*-test to evaluate

TABLE II

MEANS AND STANDARD DEVIATIONS OF CROSSMODAL AND IPSIMODAL NON-IDENTICAL REACTION-TIMES (IN M.SEC.) FOR SCHIZOPHRENICS AND NORMALS

Day and stimulus	Normals		Schizophrenics	
	M	*SD*	*M*	*SD*
Day 1, light:				
crossmodal	260	37	419	121
ipsimodal	250	53	387	118
Day 1, sound:				
crossmodal	239	53	473	214
ipsimodal	243	51	437	220
Day 2, light:				
crossmodal	252	28	403	140
ipsimodal	250	33	392	128
Day 2, sound:				
crossmodal	236	49	461	222
ipsimodal	234	41	435	209

the difference between his crossmodal and his non-identical ipsimodal reaction-times. This was done separately for reactions to sound stimuli and for reactions to light stimuli. The Mann-Whitney *U*-scores were transformed to *Z*-scores which have the following properties. *Z* close to zero indicates little or no difference between crossmodal and ipsimodal reaction-time.[5] Positive *Z* means that crossmodal reaction-time is longer, and negative *Z* means that ipsimodal reaction-time is longer. The *Z* obtained for every *S* was then used in the group comparisons between schizophrenics and normals.[6]

We computed four *Z*-scores for every *S*, one for reactions to light and one for reactions to sound on each of the two testing days. In subsequent analyses these *Z*s were combined for every *S* according to a technique described by Fisher.[7] Group means and standard deviations based on the original measures are listed in Table II.

[5] S. Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956, 123.

[6] The *Z*-score was therefore an index of the relative speed of cross- vs. ipsimodal reactions. The same *Z*-score might be obtained by two *Ss* regardless of gross differences in their speed of reacting. The use of ranking in obtaining the *Z* also minimized the effect of differences in intra-*S* variance. These safeguards were believed to be necessary to compare schizophrenics with normals since both reaction-time and intra-individual variance tend to be higher for the schizophrenics. The *Z*-score may thus be considered as a fairly uncontaminated measure of the relative lengthening of reaction-time to modality shift.

[7] R. A. Fisher, *Statistical Methods for the Behavioral Sciences*, 1938, 104.

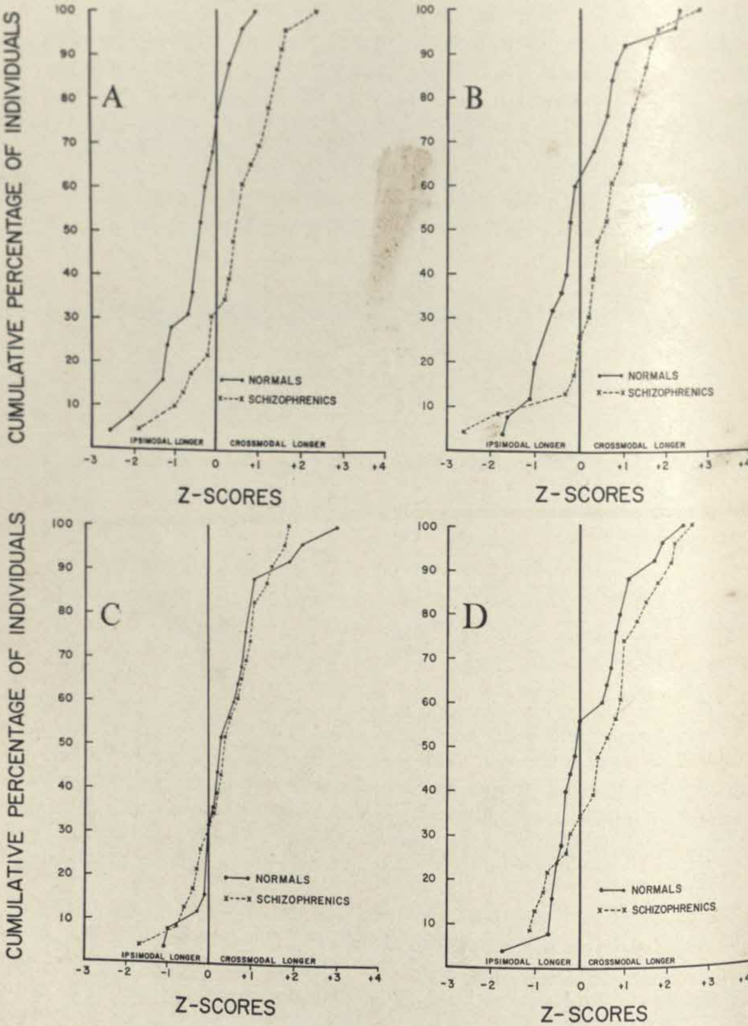


FIG. 1 A–D. RELATIVE SPEED OF CROSS- AND IPSIMODAL REACTIONS
A, to sound on first day of testing; B, to sound on second day; C, to light on first day; and D, to light on second day.

Results

The distributions of Z-scores for light and sound on two testing days are presented as cumulative curves in Fig. 1 A–D. The distributions of Z-scores for normals and schizophrenics are significantly different for sound stimuli on both test-days (Day 1, $p < 0.001$; Day 2, $p < 0.01$); but they are not different on either day for light stimuli.[8]

It will be noted that the distribution for the patients is displaced in the direction of longer crossmodal reaction-times for both light and sound stimuli on both days. The normals are less consistent. On the first testing day, they tended toward longer ipsimodal reaction-times to sound stimuli and longer crossmodal reaction-times to light stimuli. On the second test-day, however, the normals showed no particular trend toward longer or shorter crossmodal reaction-times. None of the correlations with age (rho) nor comparisons between sexes (Mann-Whitney U-test) reached statistical significance.

Discussion

The schizophrenics consistently react more slowly to a stimulus which is preceded by one in a different sensory modality than to a stimulus preceded by one in the same sensory modality. There is a reliable difference between schizophrenics and normals, however, only for reactions to sound as the second stimulus. We suspected that the lack of consistency in the comparison between the groups was due primarily to variability from trial to trial and that a larger number of trials in each of our categories would yield stable scores. We combined, therefore, all four Zs for each individual into one Z and these combined scores are presented as cumulative distributions in Fig. 2. This shows the normals to be almost symmetrically distributed around zero, *i.e.* about as many normals have longer crossmodal as ipsimodal reaction-times. The schizophrenics show only 3 individuals (13% of the sample) with longer ipsimodal reaction-times while 20 individuals (87%) have longer crossmodal reactions. The difference between the two distributions is significant at the 1% level of confidence.

It is reasonable to inquire whether the greater lengthening of the crossmodal reaction-time for schizophrenics is due to the fact that the previous stimulus was in a different sensory modality, or whether it was due to the fact that the previous stimulus was simply 'different.' It has been shown that the crossmodal reaction is longer than an ipsimodal reaction to a different stimulus. Is an ipsimodal response to a different stimulus longer

[8] Significance is estimated from the Mann-Whitney test.

than an ipsimodal response to the same stimulus? In testing for this, we also compared schizophrenics and normals as to their relative speed of reaction to identical and non-identical stimuli in the same sensory modality. Although the reaction-times for both groups to ipsimodal, non-identical stimuli were longer than the reaction-times to ipsimodal identical stimuli, no significant differences between schizophrenics and normals were found on any of the ipsimodal comparisons—all *p*s being greater than 0.05. We

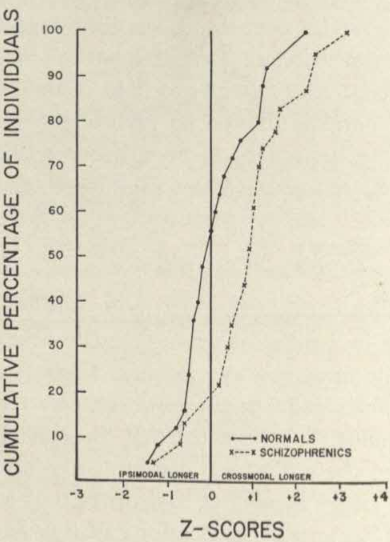


FIG. 2. COMBINED MEASURES OF RELATIVE SPEED OF CROSS- AND IPSIMODAL REACTIONS

take this to mean that the effect which is critical in discriminating schizophrenics and normals is not due to stimulus-difference *per se* but rather to differences in sensory modality.

Our experience in a large series of experiments over the last few years indicates that the lengthening of crossmodal reaction-time is not limited to schizophrenics, but also occurs in normals. Manipulation of several parameters will predictably enhance or weaken the effect. These critical parameters are the time-interval between stimuli, the number of ipsimodal stimuli which precede a crossmodal stimulus, and intensity of the four stimuli. We also find that there is a greater lengthening of crossmodal reaction to sound when it is preceded by a series of light stimuli (*e.g.* red light, red light, red light, low tone) than of crossmodal reaction to light when it is preceded by a series of sound stimuli (*e.g.* high tone, high tone, high tone, red light). Using a choice reaction-time design (a different response to each of the four

stimuli) yields similar results but at different parameter values than the design which we used in this study.

There are few direct experimental attacks on the problem of the disturbances of set and attention in schizophrenia. Rodnik and Shakow studied the ability of schizophrenics to maintain an effective set to react for varying periods of time.[9] They varied the time-interval between a ready-signal and the stimulus to which the individual was instructed to react with a simple finger movement. They found that for normal *S*s, reaction-times were consistently more rapid when trials in a given series had a foreperiod of the same duration than in a series in which the length of the foreperiod varied from one trial to the next. Schizophrenics, however, reacted more rapidly in the irregular series than in the regular series when the duration of the foreperiod was longer than 6 sec. These results were recently confirmed by Tizard and Venables.[10]

There are, however, a few studies with normals which shed some interesting light on this problem of sensory shift. Mowrer studied simple reaction-time as a function of instructional variables and the sequence of stimuli.[11] These experiments were done in the context of a controversy on central vs. peripheral determination of set. In his first experiment Mowrer found that simple reaction-time to sounds was faster when *S* was told that he would be presented with sounds only than when he was told he would be presented with both sounds and lights. This was true despite the fact that no lights were actually presented. In a second experiment, *S* was again told that both sounds and lights would be presented, but the first light was presented after a series of several sounds. The reaction-time to this 'unexpected' first light was much longer than the habitual reaction-time to light. A similar lengthening of reaction-time to an unexpected sound occurred after a series of lights. A repetition of the same experiment using two different lights (red and green) did not, however, result in a lengthening of the reaction-time to the 'unexpected' stimulus. He repeated his experiments using different permutations of sound, light, and vibration. The interesting generalization that emerges (ours, and not Mowrer's) is that reaction-time was always longer when the 'unexpected' stimulus was in a different sensory modality, and never lengthened when the 'unexpected' stimulus was in the same sensory modality. This is consonant with our finding of lengthened crossmodal reaction-time.

In another experiment, Mowrer so alternated successive stimuli that sound always followed light, and light always followed sound. After a series of 18 alternations, he 'unexpectedly' presented two sounds in a row. The reaction to the second sound was not delayed as Mowrer supposed it would be. All of these results are, however, in harmony with our findings.

We interpret the results of our experiments to mean that attention, or maximal readiness to respond (set), is not equally available to all sensory

---

[9] E. A. Rodnick and D. Shakow, Set in the schizophrenic as measured by composite reaction-time index, *Amer. J. Psychiat.*, 97, 1946, 214-225.

[10] J. Tizard and P. H. Venables, Reaction-time responses by schizophrenics, mental defectives, and normal adults, *Amer. J. Psychiat.*, 112, 1956, 803-807.

[11] O. H. Mowrer, N. N. Rayman, and E. L. Bliss, Preparatory set (expectancy)—An experimental demonstration of its 'central' locus, *J. exp. Psychol.*, 26, 1940, 357-372; O. H. Mowrer, Preparatory set (expectancy)—Further evidence of its 'central' locus, *ibid.*, 28, 1941, 116-133.

inputs at any given moment in time. The occurrence of relevant stimuli in a given sensory modality predisposes the organism to be maximally ready for further stimuli in the same modality.[12] When a relevant stimulus occurs in another sensory modality, the organism is less prepared and reaction-time is lengthened. That the central nervous system initiates some selective action with respect to the several sensory inputs would be assumed. Recent neurophysiological evidence, giving some clue as to the mechanism involved, is available in the observations of Hernández-Peón.[13] He found that evoked potentials in a sub-cortical portion of the auditory pathways of the unanesthetized cat disappear when the cat is presented with the visual stimulus of a mouse. The recent work on the role of the reticular formation in modifying *EEG* patterns may provide additional evidence to suggest the mechanism involved.

SUMMARY

Schizophrenics and normals were compared as to the relative speed of reaction to successive stimuli in the same and different modalities (vision and audition). Schizophrenics show a greater lengthening of reaction-time to a second stimulus in a different modality than do normals. A method of comparing schizophrenics and normals, so as not to yield spurious group-differences attributable to motivation, is developed.

---

[12] Relevant to the task under the instructions given. It is doubtful whether a noise outside the experimental room, unless it was of an unusual nature, had any effect on reaction-time.

[13] R. Hernández-Peón, H. Scherrer, and M. Jouvet, Modification of electric activity in cochlear nucleus during 'attention' in unanaesthetized cats, *Science*, 123, 1956, 331-332.

SEMANTIC SHIFT IN BILINGUALISM

By SUSAN M. ERVIN, University of California, Berkeley

It has long been known that languages differ widely in their systems of categories for naming colors. Roberts and Lenneberg, in contrasting the English and Zuni systems, noted that bilingual Zuni terminology differed from that of monolinguals.[1] In this paper a method of prediction of bilingual color-terminology is proposed, on the basis of a simple theory of verbal mediation. The material is presented both as a test of an extension of this theory,[2] and as an explanation for the phenomenon of semantic interference, or shift in the meaning of terms under the influence of a second language.[3]

The semantics of color must be described in probabilities, both for individual usage and in a language community. It is likely that any system of categories applied to continuous physical dimensions has this property, both because of the difficulty in learning sharp discriminations, and because the perceptual conditions for speakers and reinforcing hearers are rarely identical. At the center of categories, where the probability of a particular name is at its peak, reaction-time is consistently shorter, as is shown in comparison of Figure 1 A with 1 B. Marbe's law thus applies to naming as well as to word-associations.[4] Brown and Lenneberg have

* Received for publication April 7, 1960. This study was made possible by the sponsorship of the Southwest Project in Comparative Psycholinguistics, of the Social Science Research Council, under the direction of J. B. Carroll. The assistance of Arnold Horowitz, Herbert Landar, and Albert Silverstein, who collected the monolingual data, is gratefully acknowledged. Fig. 1 has appeared in Herbert Landar, S. M. Ervin, and A. E. Horowitz, Navaho color categories, *Language*, 36, 1960, p. 380, and is reproduced by permission of the Linguistic Society of America. The language article presents a later phonemic version of color terminology.

[1] E. H. Lenneberg and J. M. Roberts, The language of experience, *Internat. J. Amer. Linguistics*, Memoir 13, 1956, 22.

[2] J. P. Foley, Jr., and M. A. Mathews, Mediated generalization and the interpretation of verbal behavior: IV. Experimental study of the development of interlinguistic synonym gradients, *J. exp. Psychol.*, 33, 1943, 188-200; W. A. Russell and L. H. Storms, Implicit verbal chaining in paired-associate learning, *ibid.*, 49, 1955, 287-293; J. R. Bastian, Response chaining in verbal transfer, *Studies in the Role of Language in Behavior*, Technical Rep., 13, 1957, 44-45.

[3] Semantic shift has been discussed by linguists, especially Uriel Weinreich, *Languages in Contact*, Linguistic Circle of New York, 1953, 47-62; and Einar Haugen, *The Norwegian Language in America*, University of Pennsylvania Press, 1953, Vol. 2, 459-474.

[4] According to Marbe, the more common responses in word-association have shorter latencies. There is a summary of Marbe's work and later substantiations in C. E. Osgood, *Method and Theory in Experimental Psychology*, Oxford University Press, 1953, pp. 722-723.

shown that high correlations occur between measures of interpersonal agreement in naming (commonality), intrapersonal agreement, and speed of reaction, in monolinguals.[5]

Bilinguals, however, have available a larger set of possible responses. If a bilingual is asked to name colors in a particular language, he must suppress intrusions from the wrong language. If implicit responses occur in the suppressed language, response-probabilities in the overt language will be altered. Two circumstances would increase the probability of such prior implicit responses: (1) greater fluency in the suppressed language; (2) greater commonality for the primary name in the suppressed language than in the overt language. The latter condition would occur if a color were near a category-boundary in the overt language but central to a category in the suppressed language, or if a single term were dominant in the suppressed langauge for a color-range named by several terms in the overt language.

If an implicit response occurs in the suppressed language, it mediates a response in the overt language. When two responses have often been emitted in the presence of the same external stimulus, they acquire a chained relation to each other, in the sense that one later may elicit the other without the presence of the external stimulus. Synonym word-associations in monolinguals and translation-responses in bilinguals are examples. Experimentally, A-C, A-B training has been shown to produce C-B and B-C linkage.[6] The most probable terms in translation are not necessarily directionally symmetrical, since one term in one language may refer to the range of several terms in the other.[7]

It should be possible, knowing the response-probabilities of names in the two languages, a speaker's relative fluency, and translation-probabilities, to predict semantic shifts in terms.

METHOD

Subjects. On the Navaho reservation, 28 monolinguals, 21 English-dominant bilinguals, and 13 Navaho-dominant bilinguals were tested. Most of the *S*s were women, with ages ranging between 17 and 70 yr. The English monolinguals tested were 41 San Quentin prisoners, chosen for their low education. The Navahos were screened for color-blindness after naming, by having them sort chips if deviant cate-

---

[5] R. W. Brown and E. H. Lenneberg, A study in language and cognition, *J. abnorm. soc. Psychol.*, 49, 1954, 454-462.

[6] P. M. Kjeldergaard and D. L. Horton, An experimental analysis of associative factors in stimulus equivalence, response equivalence, and chaining paradigms. *Studies in Verbal Behavior,* University of Minnesota, No. 3, 1960, 21-36.

[7] S. M. Ervin, Information transmission with code translation, *J. abnorm. soc. Psychol., Suppl.,* 49, 1954, 185-192.

gories were used. The San Quentin prisoners were given the Ishihara test before color-naming.

Test of language dominance. Prior to the color-naming, the bilinguals were given a picture-test for relative speed of naming in the two languages, and those with faster responses in Navaho were designated as Navaho-dominant. This test is described in greater detail elsewhere. It has been validated by a correlation of 0.72 with years of school, the chief source of training in English, and by its significant relation to items recalled in the two languages in a recall-test. The speed obtained in the experiment reported below, when colors were named, did not correlate with the other indicators of language-dominance.

Stimulus-materials. A set of the Farnsworth Munsell 100 Hue test was prepared with acetate protectors for the color caps. These colors are designed to differ in hue but are the same in saturation and brightness. Distances are perceptually equal between the hues.[8] For testing purposes, every third chip was used.

Procedure. The colors were presented in random order, but with no colors adjacent in the hue-sequence in immediate sequence in testing. *Ss* were tested first in Navaho and later during the same day in English. They were told to talk as if they were naming the color to a friend.

Navaho instructions were tape-recorded. They were prepared by an experienced interpreter, and literally and freely back-translated before a final version was recorded. An interpreter was present to answer questions. Both the name offered and the reaction-time were recorded. If no name was given in 30 sec., the chip was re-inserted into the series and presented later.

Data-analysis. The reaction-times were converted into log. sec. and then the scores for each individual were standardized, to adjust for inter-individual differences. Names were tallied according to the head term in a construction. These were identified in Navaho by suffixes and in English either by suffix or order. Thus *blue-green* was a *green* and *greenish blue* was a *blue*. Where a pause occurred, as in *blue . . . green,* only the first term was counted. No other mergers of terms were used.

RESULTS

(1) Monolinqual. The key points of difference between the monolingual groups were first established as a basis for predicting points of semantic shift. The major differences between the two monolingual groups were as follows.

(a) Yellow. The yellow range is more difficult to name in English. The Navaho term *Litso* occurs over a wider range than *yellow* and reaches a much higher strength at it peak.[9] Color 16 was the peak in both languages, but 89% of the Navahos called it *Litso,* whereas 34% of the Anglos called

[8] Dean Farnsworth, The Farnsworth-Munsell 100 hue and dichotomous tests for color vision, *J. opt. soc. Amer.*, 33, 1943, 568-576.

[9] In transcribing the Navaho words, the following conventions have been used: 'L' refers to a voiceless lateral spirant—a breathy, whispered 'l'; 'q' refers to a glottal stop or glottalized release—the initial consonant when 'Ann' is shouted, or the consonant between the two syllables of the negative 'hunh-unh'; 'zh' refers to the final consonant of 'rouge'; and acute accents refer to raised pitch.

it *yellow*. The color is more desaturated than a good yellow, and is called *tan, beige, green,* and *brown*. English monolinguals took longer to name Color 16 than any other color except 84. Their reaction-time was con-

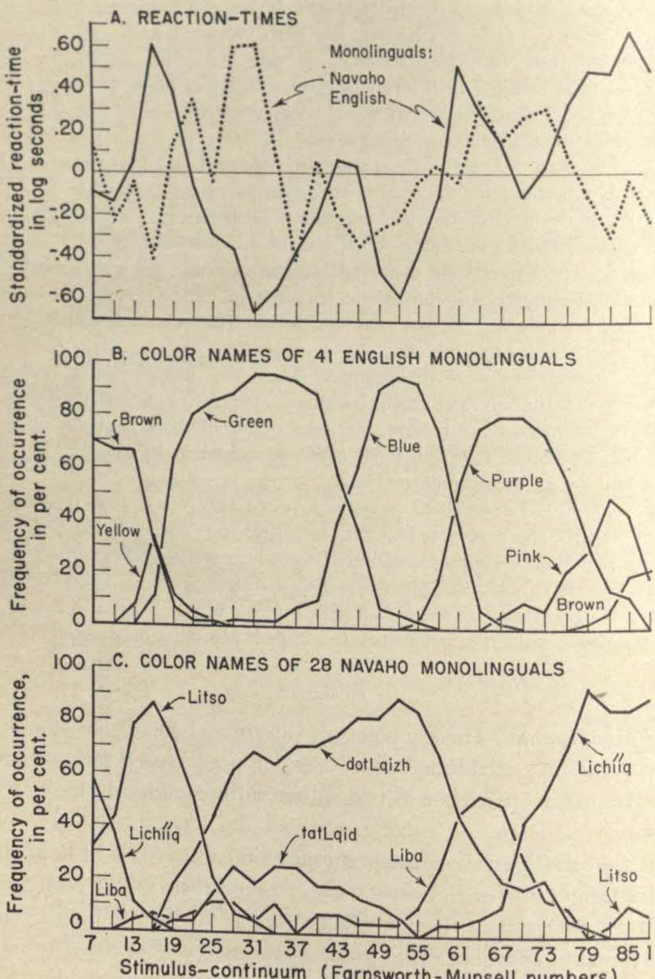


FIG. 1. DIAGRAMMATICAL DISPLAY OF SEMANTIC SHIFT IN BILINGUALISM (A) Reaction-time in log. sec.; (B) Color names of 41 English monolinguals, (C) Color names of 28 Navaho monolinguals.

siderably longer than that of the Navahos ($p < 0.0001$). On Color 16, implicit naming in Navaho would be predicted for bilinguals when speaking English, increasing the probability of *yellow*.

(b) Yellow boundary. The boundary between yellow and green is nearer green in Navaho than it is in English. For English monolinguals the transition is at Color 17; for Navaho, at Color 23.

It may be expected that at Color 17 the Navaho term will intrude for bilnguals speaking English and at Color 25 the English response will intrude at the Navaho sessions. Since the translation terms of *Litso* is less ambiguous than that for *green,* the impact of English intrusion at Color 25 should be less than that of Navaho at Color 17. Responses to the intermediate colors should be determined by the dominant language of the speaker, since these colors are near a boundary in both languages. Thus a Navaho-dominant speaker should more often have an implicit Navaho response when speaking English than an English-dominant speaker, and hence call the intermediate colors *yellow.*

The intrusion of *green* when bilinguals speak Navaho should cause maximal interference farther into the yellow range for the English-dominant than for the Navaho-dominant bilinguals. The intrusion of *Litso* at the English session should also have the effect of producing delay for the Navaho-dominant farther into the green range than for the English-dominant. In brief, the peak delay should be farther toward the green hues in the Navaho-dominant in both languages.

(c) DotLqizh. One term in Navaho, *dotLqizh,* refers to the range of three in English: *green, blue,* and *purple.* Though modifiers and some distinctive terms occur in this range, making terminological distinction possible, *dotLqizh* is the dominant term for the entire range from 25 to 67.

In English, the term *green* reaches a peak in the area of Colors 31–34, being used by 95% of the Navaho monolinguals. The corresponding Navaho term, *tatLqid* has a peak of 25% at Colors 34 and 37. The transition between *green* and *blue* occurs between Colors 44 and 45; there is no such point of shift in Navaho, since the distinctive term for green is of such low frequency. Thus the Navaho who learns English must learn a new discrimination—the transition between *green* and *blue.* It would be expected that the greater his knowledge of English, the more stable this discrimination would be, and the closer it would come to the English norm for the transition.

In Navaho, if *green* mediates his response, he may increase the frequency of its translation-terms, whatever these may be. These terms would be less likely to occur beyond Color 45, since *blue* would then mediate the response.

(*d*) *Purple.* The term *purple* has no corresponding category in Navaho. The large domain of *dotLqizh* reaches a transition to *Lichiiq* (red) at Color 69, in English the peak of the category of purple. *Purple* is bounded

on one side with *blue,* with a transition between Colors 60 and 61, and on the other side with *pink,* with the transition at Color 79.

When bilinguals speak Navaho, there should be no influence from Engish because mediation by *purple* has no unambiguous translational term. In English, however, the domain of *purple* should be reduced by the mediation of *dotLqizh* and *Lichiiq* which have other, more probable, translational terms. This effect should be stronger in Navaho-dominant bilinguals.

(e) Gray. Navahos translate *Liba* as *gray. Liba* occurred throughout the entire range of hues with low frequency, but was offered by 46% of the Navahos for Color 61. *Gray* was used rarely in English, its maximal frequency, 7%, occurring at Colors 64 and 67. If *Liba* mediates a response in English to Color 61, the frequency of *gray* should be increased in bilinguals speaking English, especially Navaho-dominant bilinguals. There would be no effect of English on Navaho, since Color 61 is at the *blue-purple* boundary.

(2) Bilingual results. The results for bilinguals are presented separately for each of the differences reported. Reported probabilities are one-tailed.

(a) Yellow. The proportion of Navaho bilinguals using *yellow* as a name for Color 16 was significantly greater ($p < 0.0005$ by X^2) than the proportion of monolinguals naming the color *yellow* ($p < 0.0005$ by X^2). The lack of competing vocabulary does not seem a suitable explanation, since the chief competing term in the English monolinguals was *brown* (32%), and this term was used elsewhere by 38% of the Navaho-dominant and 52% of the English-dominant bilinguals.

(b) Yellow boundary. At Color 24, the prediction that *green* would mediate an increase in the probability of its translation-terms in Navaho was not borne out. At Color 18 however, *Litso* and *yellow* did predominate (see Table I). For Color 21, in Navaho, but not in English, there was a significantly greater probability of using the terms for *green* (*tatLqid* and *dotLqizh)* for the English-dominant than for the Navaho-dominant bilinguals, ($p < 0.03$ by X^2). If the transitional point from yellow to green was estimated for every *S,* however, the average transition was different in English, being nearer green for the Navaho-dominant bilinguals ($p < 0.05$ by the *t*-test).

In further support of the different boundary for the two bilingual groups, in both languages the English-dominant bilinguals showed peak reaction-times farther toward the yellow direction than the Navaho-dominant bilinguals. In English this difference appeared in comparison of Colors

18 and 21, and in Navaho in Colors 21 and 24, ($p < .05$ by X^2).

(c) DotLqizh. In speaking English, the English-dominant bilinguals had more intra-individual and inter-individual agreement as to the *blue-green* boundary than the Navaho-dominant bilinguals. Five per cent of the former and 58% of the latter group showed overlap—that is, called a color *green* that was bluer than a color called *blue*. The most extreme case listed the colors from Nos. 24 to 70 as purple, purple, green, green, blue, green, green, green, green, blue, green, green, purple, green, and the rest purple. These were all hues which in Navaho could be called *dotLqizh*.

If an hypothetical transition-point for each *S* was calculated as the mid-point between the first *blue* and last *green*, the Navaho-dominant speakers had significantly greater variability, ($p < 0.01$).

There was no evidence of an increase in occurrence of *tatLqid* or of modifiers of *dotLqizh* among the bilinguals when naming colors in the green range. One reason for this is that the younger Navahos do not know the word *tatLqid* ($p < 0.005$ by the *t*-test); though learning theoretically should increase the occurrence of the translation-term, perhaps the weak initial strength of this term works against the effect.[10]

TABLE I

COLOR NAMES AND RESPONSE-TIMES OF BILINGUALS

Dominance*	Response	Color 15 Y	15 G	15 log. sec.	18 Y	18 G	18 log. sec.	21 Y	21 G	21 log. sec.	24 Y	24 G	24 log. sec.
English	English	76%	0%	−.12	62%	0%	.08	48%	14%	−.21	10%	43%	+.12
Navaho	English	69%	0%	.19	62%	8%	−.14	46%	15%	.22	8%	62%	.30
English	Navaho	95%	0%	.15	67%	14%	.02	43%	57%	.63	14%	76%	.17
Navaho	Navaho	92%	0%	.00	69%	15%	.17	69%	15%	.18	24%	54%	.76

* Among the *S*s, 21 were English dominant and 13 Navaho dominant.

Among the Navaho-dominant bilinguals, there was significantly less variability in the use of translations of *green* than among the Navaho monolinguals, ($p < 0.05$). This finding suggests that if *blue* mediates a response, the probability of *dotLqizh* (a high-probability translation) is increased, but that if *green* mediates, either *dotLqizh* or *tatLqid* may be offered.

[10] In a personal communication, Herbert Landar, who has been studying medical terms intensively, has pointed out a reason for the low frequency of *tatLqid*. The word can be analyzed at *ta-*, 'water,' and *-tLqid*, 'flatus.' Some Navahos will not use the word *tatLqid* in polite conversation for this reason, though the meaning of the combined term is 'moss, algae, or water-scum.' It is likely that the more acculturated Navahos would be more sensitive about this, hence the English-dominant would be least likely to use it.

(d) Purple. If only the responses are counted which are either *blue* or *purple,* significantly more bilinguals call Color 63 *blue* than *purple,* ($p < 0.0005$ by X^2). If *violet, lavender,* and *purple* be compared to *red* and *pink* as responses to Color 78, significantly more bilinguals chose the latter names ($p < 0.03$ by X^2). Thus for bilinguals the category of *purple* and its variant names was more limited than in monolinguals. There was no evidence that the effect was significantly stronger in the Navaho-dominant bilinguals.

(*e*) *Gray.* In the range of Colors 57 to 69, the term *gray* was not only used more often by bilinguals than by monolinguals, but significantly more Navaho-dominant than English-dominant used it at least once ($p < 0.05$ by X^2).

Discussion. Support for the treatment of bilingual naming in terms of implicit responses linked by translation to overt speech has been shown in a variety of situations. These situations can be stated in a more generalized form.

(1) In a domain where one language had a single high-probability name and the other had none, the high-probability term and its translation dominated in bilinguals in both languages (Case *a*).

(2) Where the two languages differed in the boundary between two categories, both of which have translational terms, the bilingual's dominant language determined his boundary in both languages (Case *b* and *a*).

(3) Where a category in one language covered the domain of two categories in the other language, the boundary-point in the latter language was variable and reflected the degree of learning of that language (Case *c*).

(4) A domain divided into two categories in one language was divided into three in the other, with the added category straddling the boundary in the two-category language; bilinguals reduced the size of the middle category when speaking the language with three categories (Case *d*).

We began with the assumption that the vocabulary available in the two bilingual groups would be comparable and that the difference would be due to over-all response-strength differences between Navaho and English in the two groups. This assumption was mistaken. There was evidence that English vocabulary becomes richer in color terms, with a significant relation between the language dominance score and the probability of using such terms as *lavender* and *violet,* ($p < 0.005$ by *t*-test), which have lower frequency in English than *purple.*[11]

[11] E. L. Thorndike and Irving Lorge, *A Teacher's Word Book of 30,000 Words,* 1944. Purple is rated 2 B, violet 3 A, lavender 8. Both violet and lavender were rare among the San Quentin prisoners.

In the native language of the bilinguals, the same differential appears, with the older Navahos having a richer color vocabulary. The age-difference was significant with both *tatLqid* and *Liba,* though a higher proportion of the bilinguals used the latter term. Some of the English-dominant, young, bilinguals who used these terms did so incorrectly, by monolingual standards. Thus one *S* used *tatLqid* for Color 18 and 63 but for nothing in between; therefore, the probabilities of terms as translations may not be exactly predictable from monolinguals because knowledge of the vocabulary may not be identical.

It seems likely that processes similar to those found for colors would occur in semantic shifts in other domains of meaning, such as emotion-terms. It remains to be seen whether semantic shift in the case of referents forming discrete categories, rather than referents on a continuous dimension, can be as simply explained.

These findings do not imply that there is any difference in the color vision of bilingual and monolingual *S*s at the time of immediate perception. Lenneberg's Zuni research showed no difference in Zuni and Anglo color-perception when there was simultaneous display of colors.[12] Both the arrangement of colors in order and tests of limens showed that acuity is not influenced by verbal categories. In conditions where verbal mediation may be expected to occur during performance, it is obvious, however, that bilinguals and monolinguals will differ.

SUMMARY

Semantic shift was examined in the color-naming of Navaho bilinguals, in comparison with two monolingual groups. It was found that the categories for color used by the bilinguals differed systematically from the monolingual norms. The differences could be predicted on the basis of an assumption of verbal mediation by the response-term which is most rapid. Where the response-language is restricted, the mediating term may be translated by the most probable term. The explanation accounted for four types of monolingual norm-conflict: a term of high probability for a referent in one language and not in the other; a difference in boundary between two similar categories; use of two categories in one language for the range of one category in the other; and the use of three categories in one language for the range of two in the other.

[12] E. H. Lenneberg, Color naming, color recognition, color discrimination: a re-appraisal, (in press).

THE SCALING OF PITCH BY THE METHOD OF MAGNITUDE-ESTIMATION

By Jacob Beck and William A. Shaw, University of Pennsylvania

An important problem is the degree to which ratio-scales represent stable relationships which are not affected by biases.[1] Stevens has argued that the direct estimations of sensory magnitude generally yield functions of apparent magnitude primarily determined by the functional input-output relations of the receptors,[2] but there is evidence that under some conditions response-biases and context-effects may strongly influence ratio-judgments.[3] The problem of the invariance of scale-values calls for further investigation. The purpose of the present study is to determine the effect upon the pitch-function of the frequency of the standard.

We conjectured that the frequency of the standard might affect the pitch-function because (1) the form of the pitch-function had been found to vary with the reference stimuli available, and (2) for reasons to be noted, it appeared that possible biases induced by musical relationships would differ for particular standard stimuli. Though neither consideration afforded the basis for an exact hypothesis about the shape of the pitch-function, both were in accord with the expectation that the function would be steeper with the standard set nearer the middle of the series than with a standard set at the lowest tone of the series.

The original *mel*-scale was constructed by Stevens, Volkmann, and Newman by the method of fractionation.[4] This scale was later revised by Stevens and Volkmann, who suggested that the slope of the original function was steepened because of *O*'s tendency to assume a zero-pitch that was too high.[5] For the revised *mel*-scale, the availability of a tone of 40 c.p.s., to which *O* could listen as desired, lowered the fractionational judgments. The change which occurred in the shape of the pitch-

---

* Received for prior publication December 24, 1960. This research was supported by Grant G13942 from the National Science Foundation to the University of Pennsylvania. We are indebted to Professor Francis W. Irwin for his assistance on this paper.

[1] W. S. Torgerson, *Theory and Methods of Scaling*, 1958, 421-424.

[2] S. S. Stevens, The psychophysics of sensory function, *Amer. Scientist*, 48, 1960, 226-253.

[3] W. R. Garner, Context effects and the validity of loudness scales, *J. exp. Psychol.*, 48, 1954, 218-224; Advantages of the discriminability criterion for a loudness scale, *J. acoust. Soc. Amer.*, 30, 1958, 1005-1012; The development of context effects in half-loudness judgments, *J. exp. Psychol.*, 58, 1959, 212-219; Torgerson, *op. cit.*, 421-422.

[4] S. S. Stevens, John Volkmann, and E. B. Newman, A scale for the measurement of the psychological magnitude of pitch, *J. acoust. Soc. Amer.*, 8, 1937, 185-190.

[5] S. S. Stevens and John Volkmann, The relation of pitch to frequency: A revised scale, this Journal, 53, 1940, 329-353.

function may, however, be ascribed to the presence of a reference-tone below that of any of the tones which were judged, and not to *O*'s altered conception of zero-pitch. The two interpretations are tested in the present study. Significant also in this connection is the finding of Stevens and Galanter that magnitude-estimations of pitch with no designated standard yields a function falling between the two *mel*-scales.[6]

A bias likely to assert itself in the scaling of pitch is one derived from musical relationships.[7] When the standard is the lowest frequency presented, *O* might tend to map successive octaves above the standard onto successive equal numerical intervals of the same order of magnitude, yielding a roughly logarithmic function.[8] When the standard tone is a frequency nearer the middle of the series, it is impossible for *O* (since the estimates are bounded by zero) to represent tones in octaves below the standard by mapping successive lower octaves onto successive equal numerical intervals which decrease and are of the same order of magnitude as the standard. In such a case, it seemed likely that *O* would tend to represent tones in octaves below the standard by numbers in a constant-ratio relationship. That is, a tone one octave below the standard would be assigned a number about one-half that assigned the standard; a tone two octaves below the standard would be assigned a number about one-quarter that assigned to the standard, and so forth. The expectation was that *O* would generalize and map tones above the standard onto numbers bearing similar ratio-relationships, yielding a steeper function more similar to a power-function.

In the present experiments, the pitch-function was studied under two conditions—when the standard was the lowest tone in a series and when the standard was a tone in the middle of the series. As a further check on the possibility that response-biases induced by the standard stimuli might be related to the musical relationships, the *O*s were presented (following the magnitude-estimations) with the same series of tones and required to assign numbers reflecting the octave-distance between each tone and the standard using a *preassigned numbering system*. In the first condition, the *O*s were instructed to represent pitches in successive octaves by adding successively the numerical value assigned to the standard; in the second condition, they were instructed to represent successive octaves above the standard by doubling, and successive octaves below the standard by halving. Similarity of the functions describing the magnitude-estimations and octave-relationships would not demonstrate the influence of the musical relationships, but similarity would be expected if magnitude-estimates of pitch were the result of using musical relationships and the numbering system described.

[6] S. S. Stevens and E. H. Galanter, Ratio-scales and category-scales for a dozen perceptual continua, *J. exp. Psychol.*, 54, 1957, 377-409.

[7] R. M. Warren, A basis for judgments of sensory intensity, this JOURNAL, 71, 1958, 675-687.

[8] Warren, *op. cit.*, 686.

METHOD

Observers. The total number of *O*s was 47. All were enrolled in elementary courses in psychology and served to meet course-requirements. The *O*s were summer school students who varied widely in age and background. Two experiments were conducted corresponding to the two different values at which the standard was set. Twenty-three *O*s served in Experiment I and 24 in Experiment II. Nine *O*s in Experiment I and 6 in Experiment II had served previously in a magnitude-estimation experiment. None of the remaining *O*s had served before in auditory experiments.

Apparatus and procedure. The two experiments were similar in plan. In each experiment, for each of the two different sets of instructions, 18 different frequencies were presented in a random order 5 times for a total of 180 presentations. The frequencies used were 131, 165, 196, 262, 330, 392, 440, 523, 659, 784, 880, 1047, 1319, 1568, 2093, 2637, 3136, and 4186 c.p.s.[9] The standard in Experiment I was the 131-c.p.s. tone and in Experiment II the 523-c.p.s. tone. The standard and comparison tones were presented by means of a Grason-Stadler twin oscillator (type 950-C) connected to an 8-in. permanent magnet-type loudspeaker. The sound pressure level (*SPL*) used was about 85 db. (re: 0.0002 dyne/cm.^{2}) at 6 ft. from the speaker. The oscillator and speaker were located on a desk and stool at the front of a 20 × 20-ft. classroom.

Instructions. After being seated in the classroom, the *O*s were given sheets on which had been dittoed in alphabetical order (with *A* following *Z*) five columns of 18 letters each. They then were told (a) that the purpose of the experiment was to find out how different pitches sounded to them; (b) that they would be presented on each trial with a pair of tones, a standard followed by a comparison tone; (c) that the standard tone would be assigned the number 100, and that their task was to assign a number to the comparison tone reflecting its pitch relative to that of the standard; (d) that the presentation of a pair of tones would be signalled by calling out successively the letters in the columns on the dittoed sheets; and (e) that they should base their judgments of the relative pitches on the *sounds* of the tones. In Experiment I, the *O*s were told also that all the comparison pitches would be equal to or greater than that of the standard. In Experiment II, the *O*s were told that the comparison pitches would be either equal to, less than, or greater than the pitch of the standard. To insure understanding of the instructions, examples were worked out on the blackboard together with the *O*s in Experiment I showing the numbers to be recorded if the comparison pitches were judged to be 2, 3, 4, 8, or *n* times the pitch of the standard; in Experiment II, to these same examples were added examples showing the numbers to be recorded if the comparison pitches were judged to be ½, ⅓, ¼, ⅛, or $1/n$ times the pitch of the standard. The *O*s were told further that the pitches presented were not necessarily simple ratios of the standard, and that, if they judged the pitch of the comparison to be 'almost three' or 'between four and five' times the pitch of the standard, the numbers they recorded were to reflect these judgments.

Following the magnitude-estimations, the *O*s were presented with the same series of pitches again and required to judge the octave-relationship between each pitch

---

[9] These frequencies were chosen to permit a comparison between the magnitude-estimations of pure tones and the tones of a piano. The data from the piano will be presented in a separate paper.

and the standard. Except for the instructions concerning the judgments to be made, the proceure was the same as that for the magnitude-estimations. In Experiment I, the *O*s were instructed to represent pitches in successive octaves by adding 100. Thus, if a pitch were judged to be one octave above the standard, the number 200 would be recorded; if two octaves above the standard, the number 300 would be recorded; and so forth. In Experiment II, the *O*s were instructed to represent pitches in successive octaves above the standard by multiplying by 2, and successive octaves below the standard by dividing by ½. If the comparison pitches were judged to be 1, 2, 3, 4, or *n* octaves above the standard, the numbers 200, 400, 800, 1600 or 100×2^n should be recorded; if the comparison pitches were judged to be 1, 2, 3, 4, or *n* octaves below the standard, the numbers 50, 25, 12½, 6¼ or $100 \div 2^n$ should be recorded. Examples were worked out on the blackboard until all *O*s understood the instructions. The *O*s were cautioned again that the comparison pitches did not necessarily stand in whole-octave relationships to the standard. No examples were given or questions answered concerning what was meant by an octave. The *O*s were instructed simply to do the best they could. Following their judgments, *O*s answered a short questionary concerning their musical knowledge and training.

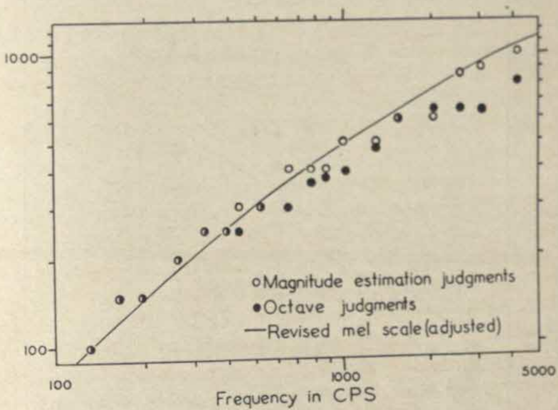


FIG. 1. MEDIAN MAGNITUDE-ESTIMATIONS AND OCTAVE-JUDGMENTS IN EXPERIMENT I
(The revised *mel*-scale also is shown, replotted with the value for 131 c.p.s. equal to 100.)

To investigate the effects of the prior magnitude-estimation judgments upon octave-judgments, two control groups ($N = 20$ in each) were run in which *O*s made only octave-judgments. The instructions and procedure for one control group were the same as those for the octave-judgments of Experiment I and for the other control group the same as those for the octave judgments of Experiment II.

RESULTS

In Figs. 1 and 2 are presented the median magnitude-estimations and judged octave-relations for Experiments I and II, together with the original

and revised *mel*-scales replotted with the values for 131 c.p.s. and 523 c.p.s. set equal to 100. Two facts are clear: (1) there are marked differences in the magnitude-estimation functions obtained in Experiments I and II accompanying the change of the standard stimulus from the lowest frequency presented to a frequency in the middle of the series; and (2) there

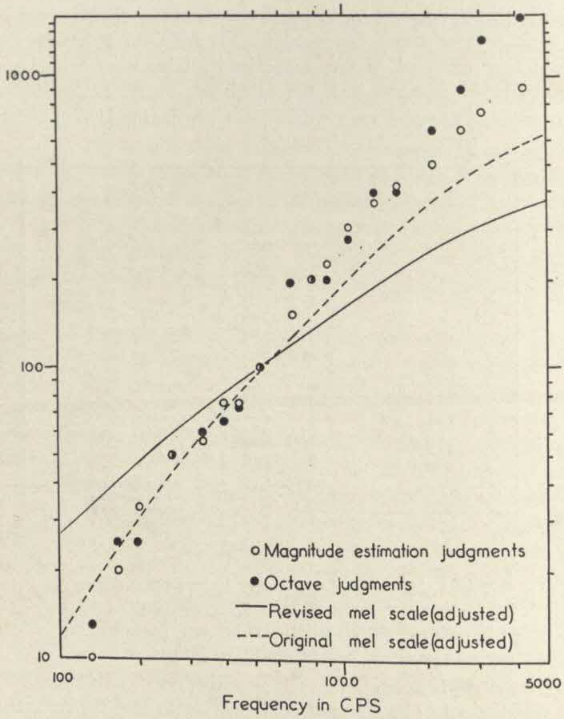


FIG. 2. MEDIAN MAGNITUDE-ESTIMATIONS AND OCTAVE-JUDGMENTS IN EXPERIMENT II
(The original and revised *mel*-scales also are shown, replotted with the value at 523 c.p.s. equal to 100.)

is a marked similarity in both experiments of the functions describing the magnitude-estimations and the octave-judgments.

The differences in the shape of the functions of the magnitude-estimations obtained in Experiments I and II reflect the tendency of *O*s to represent proportional increases of frequency by numerical differences which are more constant in Experiment I than in Experiment II. The magnitude-estimations obtained in Experiment I conform to the revised *mel*-scale. In Experiment II, the magnitude-estimates yield a much steeper function fall-

ing closer to the original *mel*-scale. In Experiment I, the magnitude-estimations above 500 c.p.s. tend to lie above the octave-judgments. For the 10 non-identical values in Fig. 1, the magnitude-estimations are above the octave-judgments nine times (although sign-tests comparing magnitude-estimations and octave-judgments at individual frequencies do not yield statistical significance). In Experiment II, the magnitude-estimations above 2000 c.p.s. fall below the octave-judgments, making the magnitude-estimations concave downward. (Sign-tests comparing magnitude-estimations and octave judgments at the two highest frequencies were significant beyond the 5% error-rate per experiment.) The results show that magnitude-estimates of pitch may be influenced by response-biases induced by the standard stimuli. The change in the shape of the magnitude-estimation functions obtained in Experiments I and II, and the similarity of the magnitude- and octave-judgments within each experiment, are consistent with the view that magnitude-estimations were affected by the musical relationships, and that the response-biases induced by the standard stimuli affected judgments by the way the musical relationships were identified with numerical relationships.

TABLE I
INTERQUARTILE RANGES FOR THE MAGNITUDE- AND OCTAVE-JUDGMENTS

Frequency	Experiment I		Experiment II	
	Magnitude	Octave	Magnitude	Octave
131	100–100	100–100	10–25	6–25
165	129–200	138–150	19–37	13–30
196	148–200	140–200	25–45	25–43
262	198–300	179–219	37–50	25–50
330	190–300	198–350	45–70	50–75
392	197–413	200–375	70–80	50–75
440	238–400	200–325	75–85	50–80
523	294–500	275–425	100–100	100–100
659	279–525	274–400	128–200	148–205
784	300–625	300–481	150–300	200–400
880	286–513	323–481	168–338	178–375
1047	300–700	300–500	200–400	200–750
1319	379–825	388–414	250–700	263–800
1568	400–800	408–713	250–600	325–800
2093	408–881	488–775	300–800	400–1200
2637	510–992	545–900	390–900	460–1600
3136	529–1025	560–963	480–1000	800–1600
4186	500–1200	650–225	495–1000	850–3200

An important question is whether *O*'s individual estimations of magnitude have the same shape as the functions graphed. Since only five judgments of a frequency were made by each *O*, the individual curves cannot be considered completely stable. The interquartile ranges presented in Table I show, however, that the variability was considerable and suggest that

caution is required in generalizing from the sample-functions to the functions of individual *O*s. In Experiments I and II, it was possible to find *O*s who represented successive octaves by approximately constant numerical differences and by increasing numerical differences. The differences in the sample curves between Experiments I and II seems, therefore, to reflect the greater frequency of one kind of curve over another due to the response bias introduced by the placement of the standard stimulus. When the standard stimulus was 131 c.p.s., the *O*s tended to represent successive octaves by more constant values than when the standard stimulus was 523 c.p.s. For example, in Experiment I, 8 *O*s represented the two lowest octave-intervals (131-262 c.p.s. and 262-523 c.p.s.) by a constant numerical value, while only one *O* in Experiment II did. Furthermore, 6 *O*s in Experiment I represented three or more of the five octave-intervals occurring in both experiments by a constant value, and for only 3 *O*s were all five of the intervals represented by numerically different values; in Experiment II, only 3 *O*s represented three of the intervals by a constant value, and for 10 *O*s all five of the intervals were represented by numerically different values.

TABLE II

MEDIAN MAGNITUDE- AND OCTAVE-JUDGMENTS FOR MUSICAL AND NON-MUSICAL SUBGROUPS

Exp.	*N*	Subgroup	Judgment	Frequency					
				131	262	523	1047	2093	4186
I	10	Music	Mag.	100	200	300	325	475	525
I	8	Non-music	Mag.	100	300	450	575	850	1200
I	10	Music	Oct.	100	200	300	400	550	700
I	8	Non-music	Oct.	100	200	450	500	925	1000
II	9	Music	Mag.	10	50	100	250	400	900
II	7	Non-music	Mag.	10	50	100	400	700	1000
II	9	Music	Oct.	13	50	100	200	800	1600
II	7	Non-music	Oct.	13	25	100	450	800	3000

On the basis of the questionary, 10 *O*s in Experiment I and 9 *O*s in Experiment II were classified as 'musical,' 8 *O*s in Experiment I and 7 *O*s in Experiment II as 'non-musical,' and the remaining as 'intermediate.' 'Musical' was defined as playing an instrument at least occasionally; 'non-musical' as not having had more than two years of musical instruction during grade school.[10] Table II shows the medians for the musical and

[10] In both Experiments I and II, five of the *O*s classified as musical played an instrument regularly, and 6 of the *O*s classified as non-musical never took any musical instruction.

non-musical subgroups in Experiments I and II. To facilitate over-all comparisons, Table II abridges the data, including only those frequencies which stood in whole-octave relationships to the standard stimuli. Inspection of the table indicates that although the estimates of the musical group (except for the octave-judgments in Experiment II) tended consistently to be below the judgments of the non-musical group, both subgroups reflected the same similarities and differences with respect to the functions describing the magnitude-estimations and octave-judgments in Experiments I and II as found in the sample data. A separate analysis of the results of the *O*s who had previously made magnitude-estimates also exhibited the trends found in the sample data.

The median octave-judgments indicate that *O*s as a group identified octave-relationships relatively acurately in Experiment I and to a less extent in Experiment II. The medians for those frequencies which stood in whole-octave relationships to the standard were in Experiment I (where *O*s were instructed to represent tones in successive octaves by adding 100), 100, 200, 300, 400, 650, and 810, and, in Experiment II (where *O*s were instructed to represent tones in octaves above or below the standard by multiplying or dividing by 2^n), 13, 50, 100, 275, 650, and 1600. The octave-judgments made by the two control groups were similar to those obtained in Experiments I and II. The medians of the corresponding frequencies for the control group in Experiment I were 100, 275, 338, 475, 800, and 912, and, for the control group in Experiment II, 12½, 38½, 100, 300, 800, and 1800.

Discussion

The results indicate the *O*s were sensitive to the response-biases induced by the standard stimulus. When the standard stimulus was the lowest frequency presented, magnitude-estimations increased much more slowly than when the standard stimulus was a frequency at the middle of the series. Response-biases have their greatest influence on behavior in situations which are ambiguous or unstructured.[11] This suggests that experiments involving magnitude-estimations of pitch present ambiguous situations. The high inter-observer variability would seem also to support this contention. How does this fit the various approaches to scaling?

The view that magnitude-estimation judgments reflect a simple correspondence between frequency and the sensation of pitch would seem to

[11] L. J. Cronbach, Response sets and test validity, *Educ. Psychol. Measmt.*, 6, 1946, 475-494; Garner reports that experiments requiring half-loudness judgments present ambiguous situations, and that response-sets appear to operate in this type of experiment (*op. cit., J. exp. Psychol.*, 58, 1959, 218).

be inadequate, since the same stimuli in Experiments I and II aroused different relative judgments of their pitches. It seems likely that *O*s are unsure of what is twice or one-half the pitch of a given tone, but when required to make such judgments identify certain stimulus-relations with numerical relations. The results indicate that musical relationships, especially the octave, may be identified with numerical relationships in making magnitude-estimations of pitch. Both Wever and Warren have, in fact, suggested that the tendency of cumulative-*jnd* and magnitude-scales of pitch to be linearly related does not reflect the subjective equality of *jnd*s but simply reflects the use of the octave-relationships to measure the progression of pitch.[12] The same suggestion can be applied to category-scales. It is necessary only to suppose that *O*s represent tones in neighboring octaves by approximately equal intervals on a numerical scale or equal category differences on a category-scale to yield the approximately logarithmic functions empirically obtained. The large individual variations in the shape of the pitch-function indicate, however, that other factors also are involved in these estimates. In general, the results may be said to be consistent with the view that magnitude-estimations of pitch are the result of *O*'s response to the demands of the situation and to the total context of stimuli presented.

The close agreement between the magnitude-estimates in Experiment I and the revised *mel*-scale suggests that the shape of the function in both cases is determined by the presence of a reference-tone which is the lowest tone heard. In Experiment II, when the standard was set in the middle of the series, making the situation similar to the original *mel*-scale, the magnitude-estimates were closer to the original *mel*-scale. Doubt is cast, therefore, on the view that the revised *mel*-scale reflects a change in *O*'s conception of zero-pitch.

The frequencies employed in the experiments have definite musical relationships to one another. Consequently, care must be taken in generalizing the results to sets of stimuli for which this is not the case. From the similarity of the results for musical and non-musical *Os,* as well as those of Experiment II and an experiment with dissonant tones,[13] it appears that the response-biases are not to be traced entirely to musical relationships. The problem is being further investigated.

[12] Warren, *op. cit.*, 686.

[13] An experiment was conducted in which individual *O*s made magnitude-estimations of 11 different tones so presented that neighboring tones produced dissonances. The standard was, as in Experiment II, a frequency from the middle of the series, and was presented before each of 5 judgment-series.

Summary

Two groups of *O*s were required to make direct numerical estimates of the magnitudes of pitches. For the first group, the standard was the lowest tone of the series, while for the second group the standard was a tone in the middle of the series. Following the magnitude-estimations, the *O*s were presented with the same series of tones and required to assign numbers reflecting the octave-relationships between each tone and the standard, according to a preassigned numbering system. The magnitude-estimations and octave-judgments within each experiment were similar, but the magnitude-estimations in experiments differed markedly. In Experiment I, the magnitude-estimations yielded a function similar to the revised *mel*-scale. In Experiment II, the magnitude-estimations yielded a much steeper function, closer to the original *mel*-scale. Individual magnitude-estimations in both experiments showed considerable variability, and caution is required in generalizing from the sample function to individual functions. The results are consistent with the view that the musical relationships influence magnitude-estimations, and that the pitch-function is the result of *O*'s response to the demands of the situation and to the total context of stimuli presented.

SOME CHARACTERISTICS OF 'AMYGDALOID HYPERPHAGIA' IN MONKEYS

By J. S. SCHWARTZBAUM, University of Wisconsin

Bilateral resections in the vicinity of the amygdaloid complex have been found to affect dietary activities. In both primates and carnivores, the amygdalectomized animals typically develop hyperphagia and, as a consequence, become obese.[1] The higher rate of weight-gain may persist for a considerable period of time after the operation.[2] The animals are also less selective about what they will eat, consuming foods and other objects that are normally rejected.[3] Both sets of effects are expressions of a more complex disturbance, as described by Klüver and Bucy, in temporal lobectomized monkeys.[4]

It is of some interest to determine whether the changes in dietary behavior reflect a more general increase in hunger or drive for food. Such an effect would correspond in its consequences to an increase in deprivation of food. It would be reflected in a disposition not only to consume more food than would normal animals, but also to respond more vigorously to food. In the present experiment, the rate of performance of a response periodically reinforced with food was studied under different conditions of deprivation and satiation.

EXPERIMENT I

Subjects. The *Ss* were eight preadolescent rhesus monkeys. As detailed elsewhere, four of the *Ss* (*AM* 397, 405, 438, 442) had received bilateral resections of the amygdaloid complex and adjacent anteromedial temporal cortex, while the others

---

* Received for publication February 11, 1960. This experiment was performed at the Institute of Living, Hartford, Connecticut, under a grant, No. M-546(C), from the U. S. Public Health Service, National Institutes of Mental Health.

[1] K. H. Pribram and Muriel Bagshaw, Further analysis of the temporal lobe syndrome utilizing frontotemporal ablations, *J. comp. Neurol.*, 99, 1953, 347-375; J. L. Fuller, H. E. Rosvold, and K. H. Pribram, The effect on affective and cognitive behavior in the dog of lesions of the pyriform-amygdala-hippocampal complex, *J. comp. physiol. Psychol.*, 50, 1957, 89-96; J. D. Green, C. D. Clemente, and J. de Groot, Rhinencephalic lesions and behavior in cats, *J. comp. Neurol.*, 108, 1957, 505-545; P. J. Morgane and A. J. Kosman, Alterations in feline behaviour following bilateral amygdalectomy, *Nature*, 180, 1957, 598-600.

[2] Morgane and Kosman, *op. cit.*, 599.

[3] Pribram and Bagshaw, *op. cit.*, 355-359; Leon Schreiner and A. Kling, Behavioral changes following rhinencephalic injury in the cat, *J. Neurophysiol.*, 6, 1953, 643-659.

[4] Heinrich Klüver and P. C. Bucy, Preliminary analysis of functions of the temporal lobes in monkeys, *Arch. neurol. Psychiat.*, 42, 1939, 979-1000.

(439, 441, 443, 447) had received an equivalent sham operation.[5] The body-weights ranged from 4.3 to 6.6 lb., with no significant difference between groups.

Procedure. All tests were carried out in a sound-insulated Skinner box which is described in the report previously cited. *S* had access in the test-chamber to a lever and a food cup. Reinforcements for bar-pressing were made available in accordance with a 2-min. 'fixed-interval' schedule. Each reinforcement consisted of a ½ gm. lab food pellet (P. J. Noyes and Co.). The test-sessions were of 60-min. duration, spaced on alternate days so as to occur three times weekly; a two-day interval between sessions overlapped the weekends. The *S*s were maintained on a daily diet of 8 to 10 Purina Lab Chow pellets and one quarter of an orange, supplemented by four unshelled peanuts on non-test days. They were fed from 2 to 3 hr. after each test and 24 hr. before the succeeding test. As a check on the maintenance-regimen, body-weights were measured before each session.

The test-conditions described above furnished a behavioral base line for the variables under study. Following the preliminary training, each *S* received a total of nine such control-sessions preoperatively and, after a two-week recovery period, the same number of sessions postoperatively. Two *S*s in each group received an additional month of tests with different amounts of reinforcement.[6] Their control-levels were reëstablished before proceeding with the present experiment. This difference in experience, which was equivalent for the two groups, did not seem to affect the results.

In the experiment proper, the *S*s were switched from a 24-hr. cycle of deprivation to one which alternated *ad libitum* feeding with prolonged deprivation of food. The *S*s were tested after 70 hr. of *ad libitum* feeding, defining 'satiation'-conditions, and after a corresponding period of food-deprivation. This meant that test-sessions were now spaced every third day. The satiation- and deprivation-conditions were alternated until five tests had been carried out under each of them, the sequence always beginning with the satiation.

During the satiation-periods, which began one hour after the preceding test, the food pans in the home cages were filled with fresh Purina Chow pellets and were replenished three times daily. The amount of food made available far exceeded the animals' eating capacity. In addition, a point was made of presenting fresh chow at least one-half hour before each satiation-test. One hour after the satiation-tests, the pans and disposal trays were cleared of all food, and the *S*s then deprived for 70 hr. until the next test. It was not, however, possible to control for copraphagia, evident in two of the amygdalectomized monkeys.

Results. Fig. 1 plots the changes in performance following prolonged deprivation of food, using the three preceding control-sessions as a base line; the values in the legend refer to the range of responses in the control-sessions. It is quite clear from these data that the amygdalectomized monkeys were far less responsive than the normal monkeys to the deprivation, although they were by no means insensitive to it. The over-all group difference, averaging across sessions, is significant by analysis of variance at well beyond the 5% level ($F = 7.25$ for 1 and 6 *df.*). The apparent interaction

[5] J. S. Schwartzbaum, Changes in reinforcing properties of stimuli following ablation of the amygdaloid complex in monkeys, *J. comp. physiol. Psychol.*, 53, 1960, 388-395.

[6] Schwartzbaum, *ibid.*, 390.

between groups and repeated test-sessions could not be verified statistically ($F < 1$ for 4 and 24 *df.*). All of the amygdalectomized monkeys showed an increase in bar-pressing, but it was not systematically related to the repeated tests.

These effects of the lesion were associated with differences in predeprivation body-weight as measured at the end of each *ad libitum* feeding. The amygdalectomized monkeys exceeded the normals in weight-gain during the initial *ad libitum* feeding period and, thereafter, maintained a higher

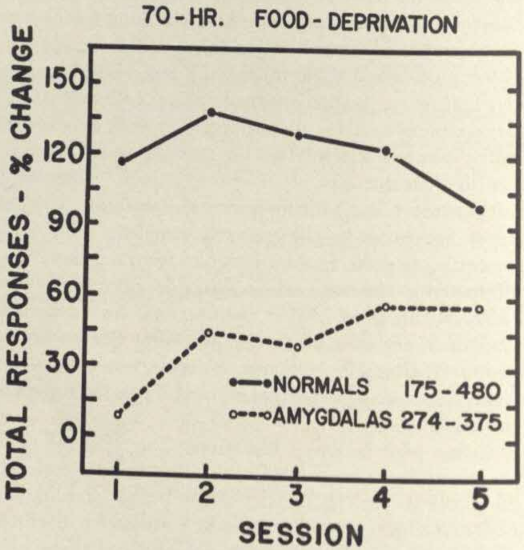


FIG. 1. MEAN PERCENTAGE-CHANGES FOLLOWING PROLONGED DEPRIVATION OF FOOD

level of body-weight both before and after deprivation. Average increases of 12% and 5% over control-levels were obtained respectively for the two groups after the *ad libitum* feedings ($p < 0.05$ by a two-tailed Mann-Whitney U test). This difference in weight-gain reflects the amount of food consumed and not the energy expended. There were no gross differences between the groups in cage activity that could account for such findings. The results, in effect, provide evidence of hyperphagic tendencies in the amygdalectomized animals.

Bar-pressing under satiation-conditions was not affected by the lesion. Both groups reduced their output by about 50% of control-levels, but variability among individual *Ss* was high.

Experiment II

In this experiment, the deprivation- and satiation-tests were separated from one another to control body-weights during predeprivation and to minimize possible interactions between the two conditions.

Subjects. A new group of five amygdalectomized monkeys (*AM* 351, 352, 395, 400, 541) and four normal monkeys (344, 390, 509, 502) were used. Each group included one cynomolgus monkey, together with rhesus monkeys. The performance of the two species was quite similar. The lesions, as reconstructed anatomically, were equivalent to those in the first experiment. The rhesus monkeys had extensive experience with discriminative problems and with bar-pressing. Body-weights of the two groups, which were not distinguishable, ranged from 5.2 to 10.0 lb.

Procedure. Test-conditions were maintained as before, except that the sessions were run daily, 6 days a week. Control-levels of performance were established under approximately 21 hr. of food-deprivation. After 10 to 12 control-sessions, the *S*s received two consecutive 70-hr. deprivation-tests with two-day intervals in the testing. Each deprivation-test was followed by a triple ration of food so as to maintain body-weights at control-levels.

A separate series of four satiation-tests was carried out after performance had been restabilized for 9 to 12 control-sessions. The *S*s were allowed to eat *ad libitum* for four consecutive days, and were tested once each day. The initial test came 22 hr. after the start of feeding. The food pans were replenished with fresh chow pellets several times a day, including once before each session. An attempt was made to obtain additional information on responsiveness to prolonged deprivation immediately after the *ad libitum* feedings. Both groups, however, showed little change in bar-pressing with respect to presatiational control-levels. Presumably, the repeated satiation-tests were responsible for this suppression, since single satiation-tests, as shown in Experiment I, did not have such a marked effect.

Results. Fig. 2 shows the intra-session patterns of performance of the two groups under control- and deprivation-conditions. It is clear from the changes in bar-pressing that amygdalectomized monkeys were again not as responsive as the normals to the increase in deprivation. An analysis was performed on the percentage-changes in total output. On the initial deprivation-test, the normal group increased its total output by 95%, in contrast to a 20% increase for the lesion group ($t = 2.48$ for 7 *df*.; $p < 0.05$). The same trend was evident on the second deprivation-test with increases of 68 and 20%, respectively, but variability among *S*s was much greater and t was not significant.

The results depicted in Fig. 2 distinguish between the effects of the lesion and increased hunger as produced by extension of the deprivation-period. Examination of the control-data shows that the lesion acted primarily to stabilize operant activity, reducing the slope of the decrement in performance. Thus, in agreement with previous findings,[7] four out of

[7] Schwartzbaum, *ibid.*, 391.

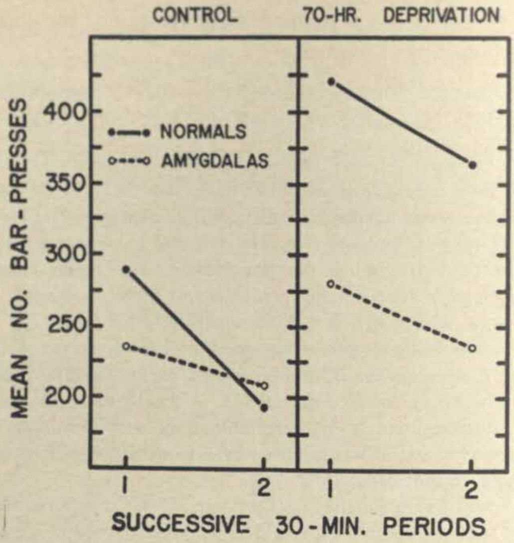


Fig. 2. Mean Number of Responses during Successive 30-Min. Periods following Prolonged Deprivation of Food

the five amygdalectomized *S*s continued to make more responses percentage-wise in the last half of the control-tests with respect to the first half than did the normal *S*s. Mean values of 93.8 and 67.8% were obtained respectively for the two groups. The group-differences in bar-presses in the first

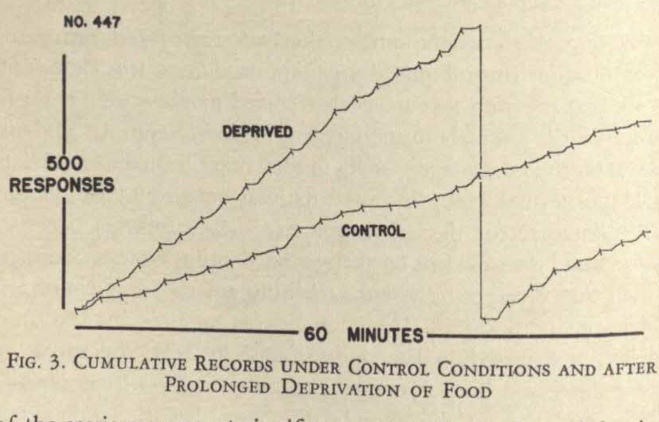


Fig. 3. Cumulative Records under Control Conditions and after Prolonged Deprivation of Food

half of the sessions were not significant statistically. Prolonged deprivation, on the other hand, increased sharply the initial or peak-rate of responding by each of the normal *S*s. It was quite clear from the individual cumulative records of performance, as illustrated in Fig. 3, that this increase occurred within the 2-min. intervals between reinforcements. Hence, while

the lesion affected the persistence of responding, the increased hunger affected the peak-rate of responding.

Bar-pressing activity after satiation, in agreement with the previous findings, was not altered by the lesion. Both groups reduced their output within the first session by approximately 68%. These results are more conclusive in the sense that the variability in performance among *Ss* was relatively low. Nevertheless, much of the bar-pressing did not seem to be motivated by the food as a 'primary reward.' About 60% of the pellets delivered were recovered after the test-sessions, with somewhat greater variability among the animals with lesions. The satiation-tests were the only ones in which this occurred. The groups again differed in weight-gain, an average increase of 15% being recorded for the amygdalectomized animals and 10% for the normal animals ($p < 0.05$ by the *t*-test). Virtually all of this increase appeared within the first 24 hr. of *ad libitum* feeding, and, again, the difference may be taken as evidence of hyperphagia.

DISCUSSION

The results indicate that amygalectomy does not increase hunger or the drive for food. First, an increase in hunger produced by deprivation influenced performance in a manner that differed from that of the lesion. Secondly, the brain-damaged animals were far less, instead of more, responsive than normal animals to prolonged deprivation of food.

A substantial increase in hunger, incident to a change in deprivation, affected the peak-rate of response within the 2-min. intervals between reinforcements. It had no consistent effect upon the persistence of responding within sessions in terms of the percentage-decrement in performance. Such changes are not specific to the conditions imposed. They vary systematically over a wide range with the degree of deprivation.[8] Thus, it is fair to assume that an effect of the lesion upon hunger, commensurate to its hyperphagic consequences, would be expressed by an increase in peak-rate of response.

Amygdalectomy did not have this effect. Although the lesion attenuated the decrement in performance within sessions, it did not alter the peak-rate or, for that matter, the total output; however, the latter might be a function of the limited duration of test-session.[9] These negative findings

[8] B. F. Skinner, *The Behavior of Organisms,* 1938, 341-405; H. F. Harlow, Primate learning, in C. P. Stone (ed.), *Comparative Psychology,* 3rd ed., 1951, 211.

[9] It should be apparent from these findings that the processes which control the peak-rate of response within the intervals between reinforcements can be isolated to a certain extent from the processes which maintain these rates across intervals. This differentiation provided by a fixed-interval schedule contrasts with that of a fixed-ratio schedule (*cf*., Murray Sidman and W. C. Stebbins, Satiation effects under

agree with pre- and postoperative comparisons of rate of performance for food, and with other operant data obtained several months after the surgery[10] More persuasive, however, is the decreased responsiveness of amygdalectomized monkeys to prolonged deprivation, since an increase in hunger would imply exactly the opposite result. Indeed, the consistently lower level of performance with prolonged deprivation suggests that under some conditions amygdalectomized monkeys may have a *lower* than normal drive for food. This effect was not likely due to some initially higher level of drive, since it occured after *ad libitum* feeding, as well as after restricted amounts of feeding. Nor can the effect be attributed to some form of debilitation or reduced level of activity. The control-data argue against this possibility, as do also the results of tests made on locomotor activity in amygdalectomized monkeys.[11]

There are two ways of interpreting the more persistent pattern of responding by the amygdalectomized monkeys. First, it can be attributed to an impairment in satiety mechanisms specific to hunger. This assumes that the normal decrement in performance related to the reinforcing events in the situation, rather than to other consequences of bar-pressing. Consistent with this reasoning, increased amounts of reward, which enhance the decrement, have been found to accentuate the stabilizing effects of the lesion.[12] The amygdalectomized monkeys showed little change in performance. Similarly, controlled amounts of prefeeding have less of a depressing effect upon the operant activity of amygdalectomized monkeys than of control monkeys.[13] The animals with lesions are not, however, insensitive to deprivation-conditions. When allowed to approach a point of satiation, as after the *ad libitum* feedings, their output fell within normal bounds. They may simply require an excess amount of food to attain this state.

Secondly, the more sustained response may be symptomatic of a general disturbance in habituation-processes, extending to other classes of stimuli beside food. Thus, amygdalectomized animals are reported to be generally more responsive, as well as persistent in responding, to objects in their

---

fixed-ratio schedules of reinforcement, *J. comp. physiol. Psychol.*, 47, 1954, 114-116). Still other processes are responsible for the temporal patterning of responses within the fixed intervals. Neither the enhanced deprivation nor the lesion altered the temporal distributions. Total output as an indicator of performance obviously confounds all of these measures.

[10] Lawrence Weiskrantz, Behavioral changes associated with ablation of the amygdaloid complex in monkeys, *J. comp. physiol. Psychol.*, 49, 1956, 386; Schwartzbaum, Response to changes in reinforcing conditions of bar-pressing after ablation of the amygdaloid complex in monkeys, *Psychol. Rep.*, 6, 1960, 215-221.

[11] J. S. Schwartzbaum, W. A. Wilson, Jr., and Rolande Morrissette, Effects of amygdalectomy on locomotor activity in monkeys, *J. comp. physiol. Psychol.*, in press.

[12] Schwartzbaum, *op. cit.*, *J. comp. physiol. Psychol.*, 391.

[13] Weiskrantz, Behavioral changes associated with ablation of the amygdaloid complex in monkeys, Unpublished Doctoral dissertation, Harvard University, 1953.

environment, the so-called "hypermetamorphotic" effect described by Klüver and Bucy.[14] They also show much less of a locomotor reaction-decrement with repeated tests in a novel situation.[15] These findings suggest a broader context for the present results. General factors of habituation are presumably involved in satiation, insofar as the satiation arises from repeated exteroceptive sensory consequences of the food aside from post-ingestional consequences.

Amygdaloid hyperphagia would, therefore, not appear to be associated with an increase in hunger-drive, but rather with some form of defect in satiation or habituation specific or not specific to food-consumption. Whether or not such a defect is a sufficient condition of the hyperphagia cannot be stated. The evidence of hyperphagic tendencies in the amygdalectomized monkeys indicates that both sets of effects were at least present concurrently.[16]

SUMMARY

Two experiments were performed in an effort to determine whether the hyperphagic effects of amygdalectomy reflect an increase in drive for food. Groups of monkeys that had received either bilateral ablation of the amygdaloid complex or an equivalent sham operation were tested in a bar-pressing situation under different conditions of deprivation and satiation. The bar-presses were reinforced with food in accordance with a fixed-interval schedule.

The two major findings were, first, that amygdalectomy decreased responsiveness to prolonged deprivation of food and, second, that it attenuated satiation-like decrements in performance within test-sessions. With prolonged *ad libitum* feeding, the amygdalectomized animals gave evidence of hyperphagia, but then performed normally for food.

It would appear that amygdaloid hyperphagia is not associated with an increase in drive for food, but rather with a defect in satiation or habituation that is either specific or not specific to food stimuli.

---

[14] Pribram and Bagshaw, *op. cit.*, 356; Klüver, and Bucy, *op. cit.*, 987.

[15] Schwartzbaum, Wilson, and Morrisette, *op. cit.*, in press.

[16] It may be noted that amygdaloid hyperphagia resembles in certain respects hypothalamic hyperphagia. Both appear to involve disturbances in satiation in the absence of any concomitant increase in drive for food, and, indeed, may be associated with a reduced level of drive (Philip Teitelbaum, Random and food-directed activity in hyperphagic and normal rats, *J. comp. physiol. Psychol.*, 50, 1957, 486-490; Philip Teitelbaum and B. A. Campbell, Ingestion patterns in hyperphagic and normal rats, *J. comp. physiol. Psychol.*, 51, 1958, 135-141). The large anatomical projection from the amygdala to the ventromedial region of the hypothalamus (W. R. Adey and M. Meyer, Hippocampal and hypothalamic connections of the temporal lobe in the monkey, *Brain*, 75, 1952, 358-384) further suggests a close functional relationship. But in view of the complexities of satiation, the apparent quantitative differences in the hyperphagia (Morgane and Kosman, *op. cit.*, 599), and the lack of any direct comparisons between the lesion-effects, any specific conclusions would be premature.

THE EFFECT OF MONAURAL AND BINAURAL TONES OF DIFFERENT INTENSITIES ON THE VISUAL PERCEPTION OF VERTICALITY

By KENNETH A. CHANDLER, Yale University

The present investigation is an outgrowth of studies concerned with the visual perception of verticality by the author in collaboration with Werner and Wapner.[1] These studies showed that consistent changes in what *Ss* consider vertical may be induced by monaural stimulation. Subsequent research showed also that the position of a luminescent rod in the dark accepted as vertical by *Ss* is also influenced by the initial position of the rod.[2] The effects of these variables and their interaction have been presented by Werner and Wapner in support of assumptions basic to a "sensory-tonic" field theory of perception.[3] The purpose of this study is to examine the relation between variations in intensity of auditory stimulation and other factors affecting the perception of verticality. This would appear to be a logical step from the previous work in view of the emphasis placed by Werner and Wapner upon the summative interaction of these factors.[4]

This investigation deals with three levels of intensity of monaural and binaural stimuli and three different starting positions.[5] In summary, the present study is concerned with (1) the effects of different intensities of monaural and binaural stimulation upon the perception of verticality, and

---

* Received for publication August 20, 1959. This study was supported by a Faculty Research Grant from the University of Delaware.

[1] Seymour Wapner, Heinz Werner, and K. A. Chandler, Experiments on sensory-tonic field theory of perception: I. Effect of extraneous stimulation on the visual perception of verticality, *J. exp. Psychol.*, 42, 1951, 341-345; Experiments on sensory-tonic field theory of perception: II. Effect of supported and unsupported tilt of the body on the visual perception of verticality, *ibid.*, 42, 1951, 346-350.

[2] Werner and Wapner, Experiments on sensory-tonic field theory of perception: IV. Effect of initial position of a rod on apparent verticality, *ibid.*, 43, 1952, 68-74.

[3] Werner and Wapner, Sensory-tonic field theory of perception. *J. Personal.*, 18, 1949, 88-107; Toward a general theory of perception, *Psychol. Rev.*, 59, 1952, 324-328.

[4] Werner and Wapner, Interaction of factors in the visual perception of verticality, in *A Program for the Study of the Basic Mechanisms Underlying Perception-Personality Relationships*, Progress Report No. 3; National Institute of Mental Health, United States Public Health Service, 1952-53.

[5] The term "intensity" used in reference to auditory stimuli will be used interchangeably with the term "loudness." Specifically, intensity is taken to mean sensation-level.

(2) a reëxamination of a method of prediction based upon the sum of factors influencing the visual perception of verticality.

APPARATUS AND PROCEDURE

Visual stimulus. All tests were carried out in a dark room. The task was the adjustment of a luminescent rod which, pivoted at its center, could be rotated in the fronto-parallel plane until it appeared vertical. This luminescent rod, 40 in. long and 1 in. wide, was 6.4 ft. directly in front of the *S* and at eye-level. The deviation of the rod from true vertical was measured in degrees from a suitably mounted 8-in. protractor. The *S* sat erect throughout all trials. His head was held in position by an adjustable head-rest attached to the chair on which he sat. Earphones for presenting the auditory stimuli were embedded in the adjustable head-rest. To reduce cues for orientation, the *S* was so seated that his feet did not touch the floor but rather were placed on a suitable foot-rest.

Auditory stimuli. The tones employed were obtained by the use of two Hewlett Packard Model 200 B audio-oscillators. These were coupled to the earphones through two Hewlett Packard Model 350 A 500-ohm attenuators. The 800~ tones were presented at two different loudness-levels of 80 db. and 90 db. above normal threshold. These levels were selected in view of the fact that the lower level is judged as half as loud as the higher level.[6] The tones from both audio-oscillators were matched, through absence of beats, for frequency. Voltage readings at the output terminals of the attenuators were 4.65 at 0 db. setting and 1.25 at 10 db. setting. Voltages at the earphones were 4.25 at 0 db. and 1.20 at 10 db. The earphones used were made by the Telephonic Corporation, Type TH 37, 300 Z. Onset of auditory stimulation was made gradually by reducing a series resistance to prevent startle, and preceded by 5 sec. the start of *S*'s performance. The tone remained on during the entire test-trial, which lasted about 30 sec. Appropriate switches enabled *E* to stimulate either ear separately or simultaneously at the levels specified.

Experimental design. Three different levels of monaural and binaural stimulation and five starting positions were studied. They were 0 db., 80 db., and 90 db. above normal threshold, and starting positions of 30° and 10° both counterclockwise and clockwise, and 0°. The conditions included asymmetrical as well as symmetrical binaural stimulation, *e.g.* 80-db. level to the left ear, and 90-db. level to the right ear. The order of presentation of the 45 conditions was randomized for each of 48 *S*s. Each *S* made two adjustments to the vertical under each of the experimental conditions. The data were analyzed as a factorial design with three axes of classification.

Subjects and procedure. *S* had his eyes covered with goggles at all times, except when adjusting the luminescent rod to vertical. *S* indicated to *E* the position of the rod and instructed *E* how to move the rod to make it look vertical. *S*s were 24 men and 24 women selected randomly from a group of 120 college freshmen and sophomores except for those who had noticeable auditory defects upon audiometric examination. Information as to the nature and purpose of the experiment was withheld until the entire group of *S*s was tested.

[6] Harvey Fletcher and W. A. Munson, Relation between loudness and masking, *J. acous. Soc. Amer.*, 9, 1937, 1-10; S. S. Stevens, The direct estimation of sensory magnitudes—loudness, this JOURNAL, 69, 1956, 1-25.

Measures employed. The position in which the rod appeared to be vertical to *S* was measured in degrees of tilt from true vertical, and are so reported here. The measure used to represent each *S*'s performance is the algebraic mean of the positions of the rod for the two trials of each test-condition. Positions clockwise of true vertical, as viewed by *S*, are designated by a plus (+) sign, and counterclockwise of true vertical by a minus (—) sign. Reliability of measurements is shown by the range of plus 0.87 to plus 0.96 for the uncorrected correlation coefficients between the first and second trials within each test condition.

Results

An analysis of variance reveals that the differences in auditory stimulation and starting position, had a significant effect upon the perception of verticality. There were also significant interactions between these variables.

TABLE I

Position of the Apparent Vertical Produced by Auditory Stimulation of the Two Ears, and Differences between Means for the Various Conditions

For each condition the first number is the sensation-level to the left ear, and the second number the sensation-level to the right ear. A plus deviation is clockwise from the true vertical

	Stimulus-conditions								
	0–90	0–80	80–90	90–0	80–0	90–80	90–90	80–80	0–0
Mean position	+.06	+.24	+.29	+.84	+.67	+.77	+.40	+.50	+.49
SD	1.62	1.82	1.64	1.85	1.57	1.84	1.65	1.66	1.74
Stimulus-Conditions 0–90	—	.18	.23*	.78†	.61†	.71†	.34†	.44†	.43†
0–80		—	.05	.60†	.43†	.54†	.16	.26*	.25*
80–90			—	.55†	.38†	.48†	.11	.21	.20
90–0				—	.17	.07	.44†	.34†	.35†
80–0					—	.10	.27*	.17	.18
90–80							.37†	.27*	.28*
90–90								.10	.09
80–80									.01

* A difference of 0.23 is significant at the 5% level of confidence.
† A difference of 0.30 is significant at the 1% level of confidence.

The mean scores for each of the nine combinations of auditory stimulation are presented in Table I. Each mean is for 10 observations by each of 48 *S*s. Differences between these means and the significance of the differences are indicated in the body of the table. These results are also presented diagrammatically in Fig. 1.

From Table I and Fig. 1, it can be seen that monaural stimulation on the left and right side induce opposite shifts in the position of the apparent vertical, when compared to the control. Stimulation on the left induces a shift of the apparent vertical to the right (clockwise) while stimulation on the right induces a shift to the left (counterclockwise). This also holds for those binaural conditions having a greater intensity of the stimulus to one side. The magnitude of the changes in the position of the apparent vertical induced by dichotic stimulation is commensurate with the differ-

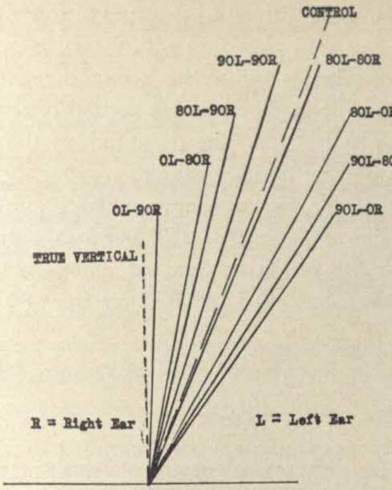


FIG. 1. POSITION OF APPARENT VERTICAL UNDER DIFFERENT INTENSITIES OF MONAURAL AND BINAURAL STIMULATION

ences in the loudness of the two tones. That is, *a tone judged twice as loud as another induces a shift in the apparent vertical which is twice that induced by the less intense tone.* This relationship is presented in Table II.

Such data suggest a linear relationship between intensity of auditory experience and the magnitude of shifts in the apparent vertical. The data presented here provide only a few points for examining this function, hence, further test of this conclusion is demanded.

Binaural stimulation with equal intensities, *i.e.* 90 L–90 R and 80 L–80 R, does not induce changes in the position of the apparent vertical from the control condition, 0 L–0 R. In other words, *symmetrically applied tones of equal intensities are equivalent in their effects to a control-condition with no experimentally produced auditory stimulation.*

TABLE II

SHIFT OF THE POSITION OF THE APPARENT VERTICAL FOLLOWING MONAURAL STIMULATION

(Tones compared which are in the relationship of 2 to 1 in loudness)

Level above normal threshold	Stimulation†	
	right ear	left ear
90 db.*	.43° CCW	.35° CW
80 db.*	.25° CCW	.18° CW

* A one-tailed test of the differences between means for conditions of 90 db. and 80 db. reveals significance at the 5% level of confidence.

† CW and CCW are clockwise and counterclockwise respectively.

Dichotic stimulation using different intensities of tones produce shifts of the apparent vertical that are in keeping with the shifts reported above for monaural stimulation. That is, with the greater intensity to the right, 80 L–90 R, there is a shift in the apparent vertical to the left of the control, and the converse is also true. The magnitude of the changes induced in the apparent vertical under these conditions are roughly comparable with the changes produced by the weaker tones applied to one ear. The average value for the second line of Table II is 0.22 and for the dichotic conditions it is 0.24. The 80 L–90 R condition does not differ significantly from the 0 L–80 R condition, nor does 90 L–80 R differ from 80 L–0 R.

It has been mentioned that the interaction between auditory stimulation and starting position is significant as tested by an analysis of variance. Table III presents the

TABLE III

THE MEAN POSITIONS OF THE APPARENT VERTICAL UNDER DIFFERENT CONDITIONS OF AUDITORY STIMULATION AND STARTING POSITIONS OF A VISUAL LINE

(Numerical values of stimulation refer to sensation-level of tones to left and right ears respectively)

Starting position	Auditory stimulation (in db.)									Mean
	90–0	90–80	80–0	80–80	0–0	90–90	80–90	0–80	0–90	
30° CW*	+2.38	+2.22	+1.68	+1.53	+1.93	+1.47	+1.22	+1.31	+ .97	+1.63
10° CW	+1.43	+1.30	+1.19	+1.19	+1.24	+1.15	+ .76	+ .79	+ .74	+1.09
0°	+ .54	+ .43	+ .48	+ .32	+ .14	+ .14	+ .12	+ .15	+ .05	+ .26
10° CCW†	− .12	+ .02	+ .32	0.0	− .12	− .24	− .07	− .24	− .27	− .08
30° CCW	− .02	− .12	− .32	− .56	− .76	− .54	− .58	− .79	−1.18	− .54
Mean	+ .84	+ .77	+ .67	+ .50	+ .49	+ .40	+ .29	+ .24	+ .06	

* Clockwise.
† Counterclockwise.

mean values for each of the five starting positions combined with each of the nine auditory conditions. The results in this Table show that when the starting position of the visual stimulus is counterclockwise of true vertical, the position judged vertical is also counterclockwise of true vertical, and the converse also holds. In addition, when the starting position of the top of the visual stimulus is on the opposite side from that of the monaural stimulation, the shift in the apparent vertical is greater than when the starting position and monaural stimulation are on the same side. This is the result to be expected from the results for each factor alone. The same relationship between these factors seems also to hold when starting position is combined with the case of dichotic stimulation.

It was shown by Werner and Wapner that it is possible to predict the position of the apparent vertical under the effect of two variables from a knowledge of the effects of each variable alone.[7] Their analysis did not, however, include differences in the intensity of auditory stimulation, although they imply that their method would be applicable to such conditions. It is evident from an inspection of Table III that, in spite of the significant interaction, the main variables account for most of the variance. On the other hand, the corners of the table show more extreme values so that there is evident some curvilinearity in the effect.

[7] Werner and Wapner, *op. cit., NIMH,* 1952-53.

DISCUSSION

It would appear from the results presented here that asymmetrically presented tones and asymmetric starting positions of the line to be judged do indeed produce consistent differences in the final position of a luminescent line which is accepted as vertical by *S*s. Tones presented at equal intensity to the two ears do not induce changes in the apparent vertical that are different from the control-condition with no tone. Moreover, it is shown that dichotic stimulation with different intensities presented to the two ears induces lesser shifts of the apparent vertical than do tones employing the more intense stimulation applied to one side alone. In general, the louder the tone the greater are the shifts in the position of what is perceived as vertical, and, within the limits of this study, a tone which is judged to be twice as loud as another, induces twice the shift of that induced by the less intense tone.

With knowledge of the effects of auditory factors and of the starting position, the simplest assumption is that the effects of these variables add linearly. Some of the evidence presented here suggests, however, that the nature of this addition is curvilinear. While the evidence reveals a significant departure from linearity it is still generally true that these factors do summate, but the function is more complex than that of a simple linear relationship.

Such findings are in keeping with a sensory-tonic theory of perception that maintains that object perception reflects the equilibrial status of the organism, and that changes in the equilibrial status of the organism are accompanied by changes in object perception.

SUMMARY

The present study was designed to investigate the effect of monaural and binaural auditory stimulation upon the visual perception of verticality. Apparent verticality was determined while *S*s were stimulated under nine different conditions of auditory stimulation and five different starting positions of the visual object. Forty-eight *S*s, 24 men and 24 women, were tested. The results are (a) that the apparent vertical was shifted away from the side of monaural stimulation and away from the side of greater intensity in dichotic stimulation; (b) that the magnitude of shifts in the apparent vertical are related in a direct fashion to the intensity of the auditory stimulus, and (c) that dichotic stimulation did not induce shifts in the apparent vertical different from those obtained with no tone, used as a control-condition. These results are in general agreement with the notion that different types of stimuli are functionally equivalent and may interact in a summative fashion.

SCALING THE ASSOCIATION BETWEEN COLORS AND MOOD-TONES

By K. WARNER SCHAIE, University of Nebraska

It is the purpose of the present study to investigate some semantic correlates of response to color. The approach to be taken will assess the strength of association between colors and mood-tones. The stability of such associations will be investigated both with respect to retest reliability within an experimental group and over different groups. An attempt will also be made to identify the dimensions underlying the system of associations by an examination of a representative sample of colors and mood-tones. Such an approach, it is hoped, will provide information regarding the methodology required for normative studies of the semantic correlates of colors and incidentally may help to generate facts and hypotheses required for a rationale for the use of the response to color as a tool for the study of personality.

Fairly complete reviews of the literature on the relation of color and affect may be found in Pressey[1] and Norman and Scott.[2] Studies directly relevant to the present problems are primarily those by Odbert, Karwoski and Eckerson[3] and Wexner;[4] a study by Tannenbaum and Osgood is also pertinent.[5] Odbert *et al.* found that certain colors were chosen more often to go with specific groups of words describing moods. A partially 'forced' method of choice, however, permitted differential emphasis on the various colors, and all moods were fitted to a color-circle (arranged according to wave-lengths) in terms of the authors' judgments.

Wexner also found relations between colors and mood-tones using more refined procedures. She prepared a list of 164 adjectives and had judges choose words which they felt to refer to mood-tones reported in the literature. Groups of words referring to 11 mood-tones were selected. Eight colors were then presented to a group

* Received for publication January 11, 1960. The author is grateful to the University Research Council of the University of Nebraska for financial assistance in carrying out this research, and to his colleagues Katherine E. Baker, Frank J. Dudek, and John C. Loehlin for their helpful criticism.

[1] S. L. Pressey, The influence of color upon mental and motor efficiency, this JOURNAL, 32, 1921, 326-356.

[2] R. D. Norman and W. A. Scott, Color and affect: A review and semantic evaluation, *J. gen. Psychol.,* 36, 1952, 185-223.

[3] H. S. Odbert, T. F. Karwoski, and A. B. Eckerson, Studies in synthetic thinking: I. musical and verbal associations of color and mood, *J. gen. Psychol.,* 26, 1942, 153-173.

[4] L. B. Wexner, The degree to which colors (hues) are associated with mood-tones, *J. appl. Psychol.,* 38, 1954, 423-435.

[5] In C. E. Osgood, G. J. Suci and P. H. Tannenbaum, *The Measurement of Meaning,* 1957, 282-283.

of students, who were asked to select the color which seemed to go best with each mood-tone, each color being permitted to be used more than once. Her results suggest that more than one color may be associated to a given mood-tone and that the same color may be associated with several mood-tones. Her method handles the latter contingency, but does not permit *S* to indicate that more than one color is felt to be strongly associated with a given mood-tone. Her results also suggest that some colors are rarely or never associated with certain mood-tones.

Osgood and Tannenbaum used the semantic differential to study color meaning. This study involved the attribution of the meaning of color in different objects. Significant differences were found to exist between colors for non-evaluative rating variables while the interactions between colors and objects were found to be significant for rating variables having evaluative properties. Factor analysis yielded significant differences among colors on factors of activity, evaluation and potency. The nature of the measurements used, however, did not provide precise measures of the relative magnitudes of the relationships between color and meaning.

PROCEDURE

A promising approach to the scaling of many psychological variables is the constant-sum method.[6] This method requires the judge to indicate the relative magnitude of two stimuli by dividing 100 points between them. The procedure generally used in the constant-sum method is patterned after the method of paired comparisons requiring a large number of indirect estimates and introducing different judgmental biases depending on the dissimilarity of the stimuli to be compared. A clearer operational meaning for the scaling operation is provided by a modified technique using the method of the constant stimulus.[7] This modification has been used as the model for the present study.

The mood-tones used in this experiment were taken from the above-mentioned study by Wexner.[8] Each of the 11 mood-tones was described by two or more adjectives as follows: exciting, stimulating; secure, comfortable; distressed, disturbed, upset; tender, soothing; protective, defending; despondent, dejected, unhappy, melancholy; calm, peaceful, serene; dignified, stately; cheerful, jovial, joyful; defiant, contrary, hostile; and powerful, strong, masterful. To get some clue as to color-preference the mood-description of 'pleasant' was added. Ten colors were used, representing the principal hues to which associations were found mentioned in the literature. Stoelting's colored papers were used; maximally saturated spectral as well as neutral colors were selected. They included: Red (Stoelting #1), Orange (#3), Yellow (#5), Green (#7), Blue (#13), Purple (#14), White (#17), Black (#18), and Gray (#19).

Since Brown is frequently mentioned in the literature but is not available in the Stoelting series, a commercial paper was used for this hue.[9] Pieces 2 × 2 in. square

[6] Milton Metfessel, A proposal for quantitative reporting of comparative judgments, *J. Psychol.*, 24, 1947, 229-235. See also A. L. Comrey, A proposed method for absolute ratio scaling, *Psychometrika*, 15, 1950, 317-325.

[7] K. E. Baker and F. J. Dudek, Scaling line-lengths with a modification of the constant-sum method, this JOURNAL, 70, 1957, 81-86.

[8] Wexner, *op. cit.*, 432-435.

[9] Approximate Munsell characteristics of the papers are as follows: Red, 5 R 4/10; orange, 10 R 5/5; yellow, 7.5 Y 8/8; green, 5 GY 8/8; blue, 7.5 PB 4/16; purple, 10 PB 3/8; white, N 9.5; gray, 7.5 PB 6/2; black, N 2; brown 10 YR 4/2.

were mounted on neutral gray backgrounds and shown in pairs on a large screen by means of an opaque projector. Red was used as the constant stimulus and paired with each of the other nine hues.

The mood-descriptions were presented in a test-booklet, all 12 moods being listed on a separate page for each of the 9 pairs of stimuli. The presentation of mood-tones was arranged in a random order which differed for each of the stimulus-pairs. Nine different page arrangements were also used to randomize the effect of a specific order of arrangement for a given stimulus-pair. The stimuli were presented in a standard-sequence following the color-spectrum and concluding with the neutral colors.

Two groups of judges were used, viz. two sections of the course in experimental psychology. The first group consisted of 23 and the second of 21 judges. Twenty-five of the judges were men and 19 were women. The judges were instructed to divide 100 points between the members of each pair of colors to indicate the relative degree of association with each of the mood-descriptions listed in the test-booklets. Prior to making the judgments the judges were informed of the implications of making various types of ratio-judgments by means of examples drawn from the judgment of line-lengths in terms of point-divisions.

RESULTS

The procedure for deriving scale-values for stimuli from point assignments has been described by Baker and Dudek.[10] The points assigned to the red stimulus in comparison with all the other hues were totaled and the ratio of the sum of all possible points to the total for red was calculated for each hue compared. These scale-values were then normalized to permit comparison between scales for the different mood-tones. Table I presents the normalized scale-values for the two groups of judges. The unit for the normalized scale-values is defined as unity divided by the square root of the sum of the raw scale-values.

Examination of Table I shows that the scale-values for those colors rated high or low with respect to any given mood-description appear to be well-replicated, while there are many exchanges in rank-order about the middle of each scale. Only 3 of the 12 scales show an exchange for the color being scaled as most strongly associated from one sample to the next, and only in one of these 3 instances is there a shift other than among closely adjacent values. None of the shifts in scale-values between the two samples is significant at the 5% level of confidence (t-test of the difference between mean point-assignments).

Table I also gives the correlation for the scale-values on each scale between the two groups of judges. These range from 0.632 to 0.946 with an average of 0.816, suggesting not only fairly good scale-reliability, but also the presence of universal color-mood associates in the population.

[10] Baker and Dudek, *op. cit.*, 81-86.

TABLE I

NORMALIZED SCALE-VALUES FOR THE RELATIVE STRENGTH OF ASSOCIATION BETWEEN COLORS AND MOOD-DESCRIPTIONS
(Obtained from two independent samples of 23 and 21 judges)

Mood-descriptions*

Color	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)		(12)	
Red	369	388	290	280	321	307	220	230	461	499	199	184	224	180	307	338	306	317	373	465	426	461	327	266
Orange	453	484	244	264	279	290	182	230	268	295	163	187	191	173	150	199	324	377	433	316	217	287	278	246
Yellow	483	540	253	252	272	238	340	251	226	191	108	147	217	194	144	152	658	520	239	344	188	150	419	358
Green	380	226	192	258	295	244	257	295	180	172	193	172	215	269	107	128	361	386	302	323	124	136	229	322
Blue	380	387	517	516	212	170	566	503	360	348	185	191	435	539	500	380	340	471	224	210	396	394	573	606
Purple	217	194	372	287	339	362	303	276	406	283	386	412	350	241	500	407	187	175	246	287	337	311	226	241
White	199	202	299	323	224	194	332	412	254	349	195	197	377	473	261	378	232	214	168	178	136	221	272	283
Gray	122	105	377	256	263	405	364	340	272	220	360	577	484	363	251	213	125	139	190	191	145	167	275	234
Black	147	149	186	224	490	485	158	190	318	409	629	458	233	231	402	501	076	097	545	470	593	552	154	131
Brown	118	106	291	393	372	328	237	307	314	241	370	304	278	265	242	240	113	107	231	212	243	175	187	245
r_{12}	.632		.769		.803		.880		.737		.716		.797		.847		.921		.825		.946		.914	

* Key to mood-descriptions: (1) exciting, stimulating; (2) secure, comfortable; (3) distressed, disturbed, upset; (4) tender, soothing; (5) protective, defending; (6) despondent, dejected, melancholy, unhappy; (7) calm, peaceful, serene; (8) dignified, stately; (9) cheerful, jovial, joyful; (10) defiant, contrary, hostile; (11) powerful, strong, masterful; (12) pleasant.

TABLE II

INTERCORRELATIONS AMONG THE SCALE-VALUES FOR THE STRENGTH OF ASSOCIATION BETWEEN COLORS AND MOOD-DESCRIPTIONS
(Above diagonal: Sample I = 23 judges; Below diagonal: Sample II = 21 judges)

Mood-Descriptions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)		−.05	−.37	.09	−.10	−.47	−.32	−.23	.60	.09	−.09	.40
(2)	.04		−.54	.86	.44	−.14	.81	.64	−.12	−.58	.05	.67
(3)	−.52	−.55		−.70	.23	.74	−.44	.19	−.49	.70	.64	−.68
(4)	−.10	.82	−.63		.03	−.33	.73	.38	.27	−.68	−.16	.81
(5)	.08	.10	.14	−.09		.22	.14	.73	.31	.16	.68	.12
(6)	−.71	−.27	.82	−.14	−.05		.10	.41	−.71	.38	.52	−.56
(7)	−.24	.60	−.46	.96	.02	.02		.49	−.36	−.65	−.14	.31
(8)	−.30	.19	.32	.09	.74	.32	.27		−.42	.02	.69	.11
(9)	.71	.17	−.59	−.15	−.13	−.63	.02	−.36		−.16	−.32	.62
(10)	.29	−.54	.41	−.78	.48	−.09	−.70	.21	.07		.66	−.36
(11)	.06	.05	.36	−.21	.87	.13	−.07	.81	−.13	.06		−.01
(12)	.46	.78	−.78	.76	−.09	−.52	.61	−.12	.59	−.37	−.10	

These values may be contrasted with the intercorrelations among scales for each single group of raters shown in Table II, which for all color-mood combinations average 0.054 and 0.079 for the two groups of judges respectively. It appears, then, that the variance common to the two groups of raters for any given scale is much in excess of the average relation among a sample of scales originating from the same judges.

Since the scales shown in Table I are ratio-scales (by definition of our scaling model) the assumptions for Pearson product-moment correlation are met, and intercorrelations can be computed among all the scales for both groups of judges. These correlations are shown in Table II, the correlations for Group I being above, and those for Group II below, the diagonal. Another test of the hypothesis that the scale-structure for the two groups of judges is identical was made by means of an adaptation of the Kolmogorov-Smirnow test.[11] The largest discrepancy fails to reach significance at the 1% level of confidence, and there is every reason to believe that differences in the magnitude of the intercorrelations are due to chance.

Interpretation of the scale-values obtained in this experiment suggests that the mood 'exciting, stimulating' has its strongest association with yellow and orange; 'secure, comfortable' with blue; 'distressed, disturbed, upset' with black; 'tender, soothing' with blue; 'protective, defending' with red; 'despondent, dejected, melancholy, unhappy' with gray and black; 'calm, peaceful, serene' with blue, white, and gray; 'dignified, stately' with purple, black, and blue; 'cheerful, jovial, joyful' with yellow; 'defiant, contrary, hostile' with black; and 'powerful, strong, masterful' with black.

A descriptive schema ordering the mood-descriptions with respect to each color is summarized in Table III. These findings confirm in part the results reported by Wexner although transpositions occur, particularly where the colors added in the present study assume a significantly high or low scale-position.

Reliability of scale-values. While there is good replication of scale-values from one sample of judges to the next, the question still remains as to how reliable our scales would be when scaling is repeated by the same judges. Such a repetition was carried out by the judges in Sample II after a 3-mo. interval. Scale-values obtained on the two occasions were correlated and Table IV lists the stability-coefficients for the repeated scaling. These were found to range from 0.78 for 'protective, defending' to 0.98 for 'pleasant' and 'exciting, stimulating.'

Factorial structure of the system of color-mood scales. A centroid factor-

---

[11] K. W. Schaie, Tests of hypotheses about differences between two intercorrelation matrices, *J. exp. Educ.*, 26, 1958, 241-245.

TABLE III

DESCRIPTIVE SCHEME FOR THE ASSOCIATION BETWEEN COLORS AND MOOD-TONES

Color	Strong association	Little or no association
Red	protective, defending; powerful, strong, masterful; (exciting, stimulating)*	Calm, peaceful, serene; tender, soothing
Orange	exciting, stimulating	calm, peaceful, serene; tender, soothing; dignified, stately
Yellow	exciting, stimulating; cheerful, jovial, joyful; pleasant	dignified, stately; despondent, dejected, melancholy, unhappy; protective, defending; powerful, strong, masterful
Green		dignified, stately; protective, defending; powerful, strong, masterful; despondent, dejected, melancholy, unhappy
Blue	pleasant; secure, comfortable; tender, soothing; (calm, peaceful, serene; exciting, stimulating)	distressed, disturbed, upset; despondent, dejected, melancholy, unhappy; defiant, contrary, hostile
Purple	dignified, stately; (despondent, dejected, melancholy, unhappy)	exciting, stimulating; cheerful, jovial, joyful
Brown	(secure, comfortable)	cheerful, jovial, joyful; defiant, contrary, hostile; exciting, stimulating; powerful, strong, masterful; pleasant
White	tender, soothing; (calm, peaceful, serene)	exciting, stimulating; despondent, dejected, melancholy, unhappy; defiant, contrary, hostile; distressed, disturbed, upset; powerful, strong, masterful
Gray	despondent, dejected, melancholy, unhappy (calm, peaceful, serene)	exciting, stimulating; defiant, contrary, hostile; powerful, strong, masterful; cheerful, jovial, joyful
Black	distressed, disturbed, upset; defiant, contrary, hostile; despondent, dejected, melancholy, unhappy; dignified, stately; powerful, strong, masterful	exciting, stimulating; secure, comfortable; tender, soothing; cheerful, jovial, joyful; calm, peaceful, serene; pleasant

* Parentheses indicate moderate associations or mood-tones whose scalar order varied between the two groups of judges.

TABLE IV

STABILITY COEFFICIENTS FOR SCALE-VALUES FROM REPEATED RATING BY THE SAME JUDGES

(N=20)*

(1)	Exciting, stimulating	.98
(2)	Secure, comfortable	.86
(3)	Distressed, disturbed, upset	.90
(4)	Tender, soothing	.95
(5)	Protective, defending	.78
(6)	Despondent, dejected, melancholy, unhappy	.86
(7)	Calm, peaceful, serene	.91
(8)	Dignified, stately	.86
(9)	Cheerful, jovial, joyful	.94
(10)	Defiant, contrary, hostile	.84
(11)	Powerful, strong, masterful	.92
(12)	Pleasant	.98

* One of the original judges was unavailable for the repeat experiment.

analysis was next performed on the correlation-matrix for Group I, leading to the extraction of six factors, four of which could be interpreted after appropriate orthogonal rotation. The second matrix was then factored and rotated to the structure suggested by the first matrix, resulting in essential replication, even though there were some minor differences. No direct comparison or exact test of significance seemed available, but the outcome of the test of the differences between the two correlation-matrices would tend to suggest that any differences between the samples in factor-loadings are predominantly a function of chance-variability.

Factor I was identified as an activity-passivity factor. The cluster at the active extreme included the mood-tones: 'exciting, stimulating'; 'cheerful, jovial'; and 'pleasant.' At the passive extreme the mood-tones were: 'distressed, disturbed'; 'despondent, dejected'; 'calm, peaceful'; 'dignified, stately.' The second factor seemed to relate to the positive or negative quality of the emotional tone of the mood-state. At its positive end it included the descriptions: 'calm, peaceful'; 'secure, comfortable'; 'tender, soothing'; 'dignified, stately'; and 'pleasant.' The mood-tones having negative qualities included: 'distressed, disturbed'; 'despondent, dejected'; and 'defiant, contrary.' Factor III was identified as a factor of strength or power. It included the scales: 'secure, comfortable'; 'protective, defending'; 'dignified, stately'; defiant, contrary, hostile'; and 'powerful, strong, masterful.' Factor IV seemed to be concerned with emotional control or mastery of the situation with 'protective, defending' at the positive and 'despondent, dejected' at the negative extreme.[12]

It should be stressed that the factors described above do not refer to a clustering of mood-tones *per se,* but rather to the clustering of color-associations to mood-tones. To clarify this matter, scale-values were summed for each color over the scales having substantial loadings on each factor. Factor I (activity-passivity) is positively associated with yellow and negatively associated with black and purple and to a lesser extent with gray and brown. Factor II (quality of emotional tone) is positively associated with blue, somewhat positively with gray, white and yellow, and negatively associated with black. Factor III (mood-strength) is positively associated with black, blue, purple, and red and negatively associated with green and yellow. Factor IV (emotional control) is positively associated with black, somewhat positively with purple, and negatively associated with yellow and green.

These findings invite comparison with Tannenbaum and Osgood's

---

[12] Contributions of these factors to the common variance are approximately as follows: Factor I, 26%; Factor II, 27%; Factor III, 21%; Factor IV, 4%.

data.[13] Our activity-passivity factor seems to match their activity-factor. In both studies yellow appears at the active and purple at the passive end. The match is not as good for some of the other colors. Our factor involving 'quality of emotional tone' would seem to correspond to Osgood's evaluative factor. Here there is a distinct discrepancy, since these authors report yellow to be at the positive extreme, while in the present study it was found to be at the negative end of the factor. The disagreement on the evaluative dimension may well be a function of the particular stimulus-objects selected for the study. Our factor of mood-strength could correspond to their potency-factor, for which ordering of colors in terms of saturation is reported, which seems to be supported by our findings.

SUMMARY

The association between 11 adjectival mood-descriptions as well as the term 'pleasant' with 10 colors was scaled by means of a variation of the constant-sum method using the constant-stimulus model. Scale-values were obtained for each color on each mood-tone from two groups of judges, one of which repeated judgments after a three-month interval. Scale-values show reasonably good replication from one group of raters to the other and for the same group over time. Low and high scale-values are well replicated but many changes of position occur for intermediate values. Some colors are found to be associated with several mood-tones and some mood-tones are associated with more than one color.

Intercorrelations of scale-values between colors and mood-tones were factored and four factors were identified. These factors were interpreted as the dimensions of activity-passivity, quality of emotional tone, mood-strength and emotional control.

Previous findings of associative relations between colors and mood-tones as well as some of their semantic dimensions are confirmed. A methodology producing reasonable stable scales is described. Findings of scalar consistencies as well as a wide range of individual differences suggest the desirability of future normative studies.

[13] Tannenbaum and Osgood, *op. cit.*, 282-283.

EFFECTS OF PRESENTATION OF PAIRED AND SINGLE-STIMULUS ON DISCRIMINATION OF LENGTH

By LEWIS P. LIPSITT and TRYGG ENGEN, Brown University

Comparisons of discriminative learning in human *S*s under conditions of paired and single-stimulus presentation suggest that the simultaneous procedure results in faster learning than the successive.[1] Method of stimulus-presentation is pertinent to the areas both of psychophysics and of discriminative learning. Psychophysics investigates the effect of stimulus-conditions upon discrimination. In turn, psychophysical data are basic to learning-processes by which differential responses to stimulus-differences are acquired.

Two experiments are reported here in which pairs of lines differing in length are presented both simultaneously and successively. They differ only in the nature of the response required; the first study merely requires discrimination, while the second requires the learning of differential responses. Presumably, any differences in accuracy of performance will have resulted from the different response-requirements.

THE PSYCHOPHYSICAL EXPERIMENT

In the first experiment, a standard psychophysical method was employed.

Method. White lines on a gray background were shown on a screen (by means of a 35-mm. projector) to several groups separately. The lines were presented according to the method of constant stimuli. The standard line was 66.0 mm., and the 10 comparison-lines ranged from 49.5–82.5 mm. in 3.3-mm. steps. Each comparison-line was paired with the standard 10 times, and the pairs were presented in random sequences.

Judgments were obtained from five groups ($N = 14$, 15, 16, 18, and 19) of students from elementary psychology. The *S*s had no previous experience in such experiments. The stimulus-lines were presented to each group separately. The *S*s indicated their judgments after the presentation of each stimulus-pair by marking prepared answer-sheets.

Group *P* (paired presentation) was required to judge which of two lines on a horizontal plane was the longer—the one on the left or the right. Groups S_S and S_L

* Received for publication January 25, 1960.

[1] M. T. Erickson and L. P. Lipsitt, Effects of delayed reward on simultaneous and successive discrimination learning in children, *J. comp. physiol. Psychol.*, 53, 1960, 256-260; H. B. Loess and C. P. Duncan, Human discrimination learning with simultaneous and successive presentation of stimuli, *J. exp. Psychol.*, 44, 1952, 215-221; M. J. Perkins, H. P. Banks, and A. D. Calvin, The effect of delay on simultaneous and successive discrimination in children, *ibid.*, 48, 1954, 416-418.

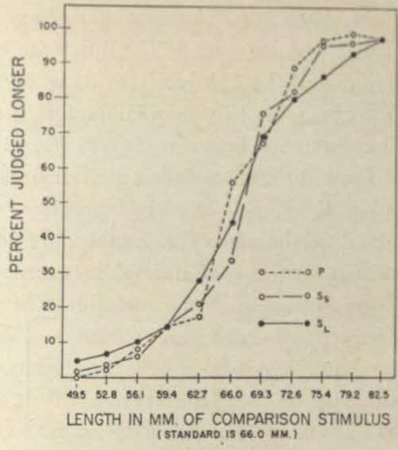


FIG. 1. MEAN PERCENTAGE OF 'LONGER' JUDGMENTS AS A FUNCTION OF LENGTH OF THE 10 COMPARISON-STIMULI FOR EACH GROUP

(single presentation) were shown the same line-lengths, each line being presented on a separate slide, and asked to judge whether the first or second line was longer. For Group S_S, the slides of each pair were projected with a 1-sec. interval between them, while for Group S_L a 5-sec. interval was used. Time between trials (pairs) was the same for all groups, approximately 8 sec.

Results. Fig. 1 presents the psychophysical function obtained under the three stimulus-conditions. The standard stimulus is indicated as 66.0 mm. on the abscissa, and mean judgment of 'longer than' for each comparison-stimulus is plotted on the ordinate. The slope of least steepness is that of Group S_L, indicating poorer discrimination under this condition. Groups P and S_S do not seem to be widely separated.

The *DL*s for each group estimated from these data by the normal graphic process are presented in Table I. This analysis suggests that paired

TABLE I

RESULTS OF THE PSYCHOPHYSICAL EXPERIMENT

	Groups		
	Simultaneous	Successive-short	Successive-long
DL (in mm.)	3.8	4.8	5.6
Mean error	9.06	9.53	13.67
SD	4.72	5.75	3.51

presentation provides somewhat finer discriminations than the single presentation, and that the longer interstimulus-interval in successive presentation results in the poorest discrimination. The data are not, however, well represented by straight lines. To evaluate the reliability of the differences

among groups, the mean number of judgmental errors for each group, also presented in Table I, was computed. The difference between Group *P* and Group S_L is significant at the 1% level ($t = 3.03$, $df. = 31$), and the difference between Groups S_S and S_L is significant at the 5% level ($t = 2.11$, $df. = 32$). The difference between Groups *P* and S_S does not approach significance. These differences reflect the same rank order of discriminability as does Fig. 1.

The results suggest that the ability to judge differences in length is rather more dependent upon the time-interval between stimulus-presentations than on whether the stimuli are presented strictly simultaneously or singly.

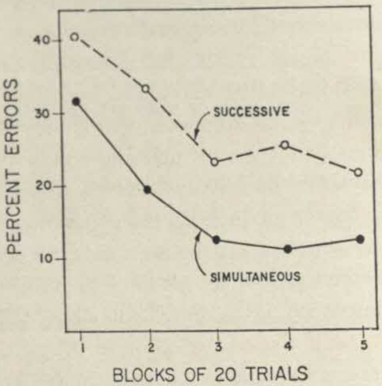


FIG. 2. MEAN PERCENTAGE ERROR AS A FUNCTION OF LEARNING TRIALS FOR THE SIMULTANEOUS AND SUCCESSIVE GROUPS

THE LEARNING EXPERIMENT

The fact having been established that the difference between simultaneous and successive discrimination of live-length is negligible if the time-interval between successive presentations is small, the same stimuli were used in an experiment with two new groups, one comparable to Group *P* and another to Group S_S of the psychophysical experiment. Now, however, a discriminative learning factor was introduced.

Method. The general conditions of the second experiment were the same as those of the first. The *Ss* were instructed to indicate after each trial which of the two stimuli was 'correct.' Following the response, *E* informed the group which stimulus was in fact 'correct'; namely, the longer of the two lines.

Results. The learning-curves for the two groups are shown in Fig. 2. It appears that the paired group does relatively better than the single group and that the advantage of this condition is maintained throughout 100 trials. Analysis of variance confirms these observations. The difference be-

tween mean number of errors for the two groups is significant at the 2.5% level ($F = 5.74$, $df. = 1, 28$), and no trials-groups interaction is found. A significant trials-effect significant at the 0.1% level ($F = 5.72$, $df. = 4, 112$) indicates that both groups learned their respective discriminations.

DISCUSSION

Since, in the psychophysical experiment, the simultaneous procedure yielded approximately the same number of judgmental errors as the successive, short-interval procedure, the difference between the simultaneous and successive, short-interval groups which appeared in the learning-experiment cannot be traced to a strictly perceptual impairment. One possible interpretation of the difference is that, under learning-conditions, *S*s have a greater opportunity to develop erroneous hypotheses or mediating responses concerning the correct response. *S*s are more likely to develop such extraneous hypotheses under the simultaneous condition than the sucessive, since twice as many slides or stimulus-settings are involved in this condition. It was noted that *S*s in the learning-experiment tended more under the simultaneous condition than the successive to select irrelevant cues, such as slight imperfections in the shading or background of the slides. There would, of course, be no such problems in the psychophysical experiment, for the instructions directed attention to the pertinent cue. It may be supposed that erroneous hypotheses would be most likely to occur in situations, such as the present one, involving highly similar stimuli.[2]

SUMMARY

Two experiments were performed in which *S*s were to determine which of two lines, presented either simultaneously or successively, was the longer. In the first, psychophysical experiment, *S*s were simply instructed to select the longer of the two lines. In the second, learning-experiment, *S*s were instructed to discover the correct stimulus of each pair; they were informed after each judgment as to which stimulus was in fact correct, but were not told that the relevant feature was length. In the psychophysical task, a simultaenous and a successive, short- interval (1-sec. interstimulus-interval) group did not differ, but the simultaneous group was significantly superior to a successive, long-interval group (5 sec.). In the learning-experiment, the simultaneous and successive, short-interval groups differed significantly, indicating that the addition of a learning-requirement to a perceptual task results in differential impairment for simultaneous and successive groups, with successive being more greatly affected.

[2] For relevant data at the level of the rat, see E. F. MacCaslin, Successive and simultaneous discrimination as a function of stimulus-similarity, this JOURNAL, 67, 1954, 308-314.

THE RATE OF PECKING BY PIGEONS AS A FUNCTION OF AMOUNT AND FREQUENCY OF REINFORCEMENT

By BERNARD PYRON and L. B. WYCKOFF, University of Wisconsin

Behavior often occurs in a situation in which not all responses by an organism are reinforced, and the frequency of reinforcement can vary widely. In addition, the amount of the reinforcing agent may vary. For example, an organism might receive 2 units of amount 40% of the time or 4 units 20% of the time. In each of these cases, over a long period, the organism receives the same total amount of reinforcement. An amount of 2 units multiplied by a relative frequency of 0.4 yields an average gain of 0.8 unit per trial and 4 units multiplied by a relative frequency of 0.2 yields the same average gain per trial. The total amount of reinforcement is essentially the same concept as the Expected Value (*EV*) of a bet, used by mathematicians to describe the product of the amount and relative frequency of reward in a gambling situation.

By systematically varying frequencies and amounts of reward we can create combinations which have the same *EV*. We can, moreover, construct a situation in which an organism may choose between two or more combinations of amounts and frequencies of reward. When *EV*s are equal, will the organism choose one frequency and amount combination as often as the other as predicted by the *EV*-model, or will he choose the combination paying off with the greater frequency or with the greater amount? We can now directly compare frequency vs. amount of reward as determinants of choice behavior.

Some experiments suggest that when the amount of reinforcement measured in terms of the concentration of a desired substance is varied, rate of response will increase logarithmically with increases in concentration.[1] Sheffield, among many others, found that 100% reinforcement yielded faster learning than 50% reinforcement.[2] Pyron, using human *S*s, found

* Received for publication June 18, 1959.

[1] Norman Guttman, Operant conditioning, extinction and periodic reinforcement in relation to concentration of sucrose used as reinforcing agent. *J. exp. Psychol.*, 46, 1953, 213-224; Guttman, Equal-reinforcement values for sucrose and glucose solutions compared with equal-sweetness values, *J. comp. physiol. Psychol.*, 47, 1954, 358-361.

[2] V. F. Sheffield, Extinction as a function of partial reinforcement and distribution of practice, *J. exp. Psychol.*, 39, 1949, 511-526.

that rate of response was a linear function of the frequency of reward and that amount of reward (poker chips) was of little importance.[3]

The present experiment was designed to vary both the amount and frequency while holding *EV* constant.

Method. Two *EVs* were used (1.2 and 2.4). Four frequencies of reward were used, 10%, 20%, 40%, and 80%. The amount was varied by the use of three lengths of time the *Ss* were allowed to eat during each reinforcement. These periods were 3 sec., 6 sec., and 12 sec.

Apparatus. A Skinner-box suitable for training pigeons was equipped with two colored lights placed behind glass disks 1.5 in. in diameter. The two lights were 3.0 in. apart. Each disk served as a key and was so connected to counters that the number of responses to each light was recorded separately. A manually operated food chamber was located below the two lights.

TABLE I

MEAN RATIO-SCORES FOR ALL *Ss*

Amount of reinforcement	Frequency of reinforcement			
	10%	20%	40%	80%
3 sec.	—	—	2.24	2.93
6 sec.	—	1.60	2.20	—
12 sec.	.95	1.46	—	—

Subjects. Four pigeons were used as *Ss*. Every bird was run under six different conditions. Two birds were run in four replications of each of the six conditions and two in only two replications of each of the six conditions. Thus, there were 12 replications of each of the six conditions.

Design and procedure. To provide a baseline for rate of responding and to correct for variations in rate of pecking from day to day, a standard condition was always compared with one of the six variable conditions. The standard condition required a response to one light (either red or green) and the variable a response to the other. Both lights were on at the same time. The standard condition throughout the experiment was one reinforcement in 10 trials with the food chamber open for 3 sec.

The six variable conditions included three combinations having an *EV* of 1.2 sec. of reinforcement per trial. The relative frequencies and amounts were as follows: 10% of the trials with the food available for 12 sec., 20% and 6 sec., and 40% and 3 sec. The three combinations having an *EV* of 2.4 were 20% and 12 sec., 40% and 6 sec., and 80% and 3 sec. The six combinations of frequency and amount are shown in Table I.

To prevent the *Ss* from learning a constant discrimination based on the lights alone, one light was used for the variable condition for a block of 120 trials, and, for the

[3] Bernard Pyron, An attempt to establish empirical indifference curves. Unpublished M. S. thesis, University of Wisconsin, 1958.

next block of 120 trials, the other light represented the variable condition. When all six variable conditions had been run for an *S*, a different order of these conditions was substituted for the next replication.

A trial consisted of 30 sec. during which the lights were on and the *S*s responded, followed by 15 sec. during which the lights were off. If the schedule called for a reinforcement, it followed the first response to the appropriate key during the last 10 sec. that the lights were on. In this way it was possible to obtain a measure of rate of responding on each key for every trial. Time required for eating was restricted to the interval in which responses were not recorded. This procedure resulted in an average rate of response to the variable during a block of 10 trials of 204.7 pecks. Thus, with rare exceptions, the *S*s received the prescribed number of reinforcements in 10 trials.

RESULTS

The data were analyzed only for the last 100 trials in a block of 120 trials. The first 20 trials were considered a transitional period. The mean number of pecks for 10 trials was computed for both the standard and the variable. A ratio-score was obtained for each block of 10 trials by dividing the number of responses to the standard into the number of responses to the variable. The mean of the ratio-scores was then obtained for each 100 trials of a variable condition.

The mean ratio-score for each of the six variable conditions is reported in Table I. For each *EV* an increase in relative frequency results in an increase in the mean ratio-score. The same results are shown in Fig. 1 in which the scores are plotted as a function of $\log_2$ of the number of reinforcements in 10 trials. The score is almost a linear function of the logarithm of the frequency. In contrast, as can be seen from Table I, the amount of reinforcement as we measured it had little effect on the ratio-scores. An increase from 6 to 12 sec. in the time the food was available resulted, for 20% reinforcement, in a slight *drop* in the mean score. The same is true for 40% reinforcement. Statistically, the difference between the mean scores of the two *EV*s is not significant.

An analysis of variance, including orthogonal polynomials was performed on the three conditions of each *EV*. For the *EV* of 1.2, the difference between the scores for the three levels of reinforcement (10%, 20%, and 40%) was significant at the 1% level. The linear term was also significant at the 1% level. For the *EV* of 2.4, the difference between the scores for the three levels of reinforcement was not significant nor did the linear term reach significance. This was obviously due to greater variability in these conditions. For both *EV*s the quadratic term was almost zero.

Since only 4 *S*s were used, the differences between them may be of interest. In general, the same linear relationship between relative frequency of reward and scores holds for all *S*s. The mean ratio-score for each *S* in all of the six conditions was

computed. An analysis of variance was performed including individual *Ss* as a variable. The *Ss* were not significantly different.

Discussion

This experiment suggests that the rate of pecking by a pigeon is not significantly influenced by the length of time the food-chamber is open.

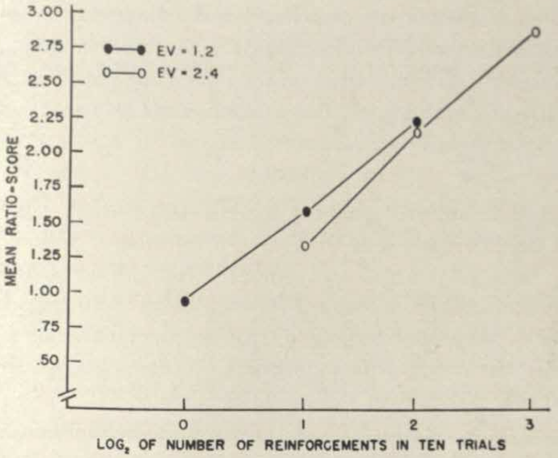


FIG. 1. MEAN RATE OF PECKING
The ratio is shown as a function of the logarithm of the number of reinforcements for *EV*s of 1.2 and 2.4.

Other investigators have found that increases in amount of reinforcement did not lead to increases in behavioral measures such as speed of running or rate of responding. Kling, for example, varied amount of reinforcement in two ways, by varying the time *S* was in the goal-box where food was obtainable and by varying the size of the drinking tube.[4] He measured speed of running in a straight alley after rats had been deprived of water for 20 hr. Neither the size of drinking tube nor duration influenced running speed. The most rapid running occurred in the group receiving the least water for the shortest duration.

Summary

Both the relative frequency of reward and the length of time the rewarding food chamber was open were varied independently in an experiment us-

[4] J. W. Kling, Speed of running as a function of goal-box behavior, *J. comp. physiol. Psychol.*, 49, 1956, 474-476.

ing the rate of pecking as an operant response. Applying the concept of Expected Value (*EV*), we compared directly frequency vs. amount of reward as determinants of choice behavior. Pigeons were run in a Skinner-box in which choices were made between a light that paid off with a constant amount and frequency of reward and another light that paid off with one of six variable combinations of relative frequency and amount.

It was found that the rate of pecking was a linear function of the logarithm of the number of reinforcements in 10 trials, and that the duration of the reward had little effect on the rate. Thus, the choice behavior of our *S*s deviated from that predicted by the objective *EV* model.

RETROACTIVE INHIBITION IN INCIDENTAL LEARNING

By SHELDON ROSENBERG, The Training School, Vineland, New Jersey

Studies by Gleitman and Kamrin, Postman and Adams, and Prentice strongly suggest that retroactive inhibition (*RI*) not only occurs for materials which *S*s are specifically instructed to learn, but for incidentally learned items as well.[1] These studies are similar in that they all used essentially the same type of procedure for creating conditions of incidental learning. Two groups of *S*s are exposed to the same set of stimulus-materials, of which one group is instructed to learn while the other is not. The procedure for orienting incidental *S*s to the learning materials can never be an instruction to learn in this design.

Although this is a common procedure in research on incidental learning, an alternative method has been employed in which all *S*s are presented simultaneously with two different classes of materials, instructed to learn only one of them, but later tested for retention of both.[2] Although this procedure does not easily permit the direct comparison of intentional and incidental learning of the *same* material, it is useful in studying the effects of various independent variables upon both varieties of learning, and represents an important everyday instance of incidental learning. Bruner has referred to this phenomenon by the term, "breadth of learning,"[3] and in animal experimentation it is somewhat similar to the variety of latent learning known as "irrelevant-incentive learning."[4] The present experiment is a study of *RI* in incidental learning using this alternative procedure, to determine the generality of results obtained with the traditional design.

---

* Received for publication, November 20, 1959. These data were collected while the author was on the faculty of Jeffersonville Center, Indiana University. The assistance of Mr. Robert Detamore, and of the administration of the Jeffersonville High School in obtaining *S*s, is gratefully acknowledged.

[1] Henry Gleitman and Robert Kamrin, Proactive and retroactive inhibition in intentional and incidental learning, *Psychol. Rep.*, 3, 1957, 155-160; Leo Postman and P. A. Adams, Studies in incidental learning: III. Interserial interference, *J. exp. Psychol.*, 51, 1956, 323-328; W. C. H. Prentice, Retroactive inhibition and the motivation of learning, this JOURNAL, 56, 1943, 283-292.

[2] H. P. Bahrick, Incidental learning under two incentive conditions, *J. exp. Psychol.*, 47, 1954, 170-172; Sheldon Rosenberg, Exposure interval in incidental learning, *Psychol. Rep.*, 5, 1959, 675.

[3] J. S. Bruner, Jean Matter, and M. L. Papanek, Breadth of learning as a function of drive level and mechanization, *Psychol. Rev.*, 62, 1955, 1-10.

[4] Donald Thistlethwaite, A critical review of latent learning and related experiments, *Psychol. Bull.*, 48, 1951, 97-129.

Method

Subjects. The *Ss* were college freshmen and senior high-school students in courses in introductory psychology. There were an Experimental Group and a Control Group of 20 *Ss* each. The *Ss* were assigned to the two conditions as they appeared for the experiment according to a schedule devised from a table of random numbers.

Materials. Materials used in original learning (*OL*) consisted of a series of 12 names of common objects (List A), *e.g. radio, dish, arrow,* etc., each of which was accompanied by a different two-digit number. The numbers were designated as incidental items. In an attempt to maximize retroactive inhibition (*RI*), materials used in interpolated learning (*IL*) (List B) also consisted of a series of 12 names of common objects. No incidental items, however, accompanied the words of List B.

Every item was printed in black ink in the center of a 4 × 5-in. white card. Individual letters and numbers were approximately ⅜-in. high and ¼-in. wide. The incidental numbers in List A appeared directly to the left of the word.

The cards were exposed manually by *E* in unison with the click of a metronome. An exposure-interval of 2 sec. was used. The order of presentation of Lists A and B was the same on all trials for all *Ss*. The cards were exposed one after another with no intervening blank intervals.

Original learning. The *Ss* were tested individually in a quiet room. List A was presented to all *Ss* for three trials, and performance after Trial 3 was taken as the measure of the degree of *OL*. Instructions for initial learning, which took 15 sec. to present, were to try to remember as many of the words as possible without regard for order. No mention was made of the incidental numbers. The *Ss* did not pronounce the words during the exposure-trials. The interval between each presentation of List A and the recall test was 7 sec., during which *S* was given a sheet of blank paper and instructed to write down, in any order, as many of the words as he could remember. Guessing was urged, and *S*'s score was the total number of correct words recalled. The time-limit for recall was 2 min. Trials 2 and 3 were presented 6 sec. after completion of the recall-test for the previous trial. During the 6-sec. interval the initial learning instructions were briefly repeated. The *Ss* were given no indication during *OL* of the fact that they would later be tested again for retention of List A.

Interpolated learning. Interpolated learning (*IL*) was begun 15 sec. after the last test of *OL*. This interval was used to present learning instructions similar to the initial instructions of *OL*. Unlike List A, however, List B contained no incidental items. The procedure for *IL*, with respect to sequence and timing, followed exactly that for *OL*. Following *OL*, the Control Group was given a simple puzzle to work on for a period of 7 min. Work on the puzzle was preceded by 30 sec. of instruction and followed by 30 sec. of discussion of the solution. The total length of the rest-interval for the Control Group, then, was 8 min. The length of the rest-interval was determined in preliminary research which revealed that with 27 sec. required for presentation of the learning instructions, 72 sec. for three presentations of List B, 21 sec. for free recall instructions, and 6 min. for three recall-tests, the total time required for *IL* would be 8 min.

Recall. Thirty sec. after the end of *IL*, both groups were given a 2-min. written test of free recall for the words in List A, which was followed 40 sec. later by a test of recognition for the 12 two-digit incidental numbers. The intervals of 30 and

40 sec. were used for the presentation of the appropriate instructions. The test of recognition consisted of a matrix of 36 two-digit numbers ranging from 11 to 46, which contained in random order the 12 incidental items. The 36 numbers were arranged in numerical order in two columns of 18 each. Every *S* was given a copy of the matrix and instructed to circle the numbers that had accompanied the words of List A. All *S*s were required to make 12 choices even if they were not equally certain of each choice. No time-limit was set for this test.

At the end of the experiment, all *S*s were told that they had performed well. An informal inquiry revealed that none of the *S*s had attempted to learn the incidental numbers, or had been aware that they would later be tested again for retention of List A.

RESULTS

The results are summarized in Table I. The difference between the means of the Experimental and Control Groups in *OL* was not significant ($t = 0.63$, $p > 0.05$), *i.e.* there was no significant difference in the amount of intentional learning. Table I also shows that recall of List A following *IL* was poorer in the Experimental Group than in the Control Group. The variances for the Experimental Group (8.83) and the Control Group (3.75) were not homogeneous ($F = 2.35$, $p < 0.05$). To evaluate the mean differences, a normal-approximation sum-of-ranks test corrected for continuity and tied ranks was used.[5] The value of z for this comparison was 4.41 ($p < 0.001$). Retention for the intentionally learned List A was significantly poorer under the experimental condition than the control condition. This finding represents, of course, the expected *RI* effect.

TABLE I

MEAN SCORES: LEARNING AND RETENTION

Group	Intentional items				Incidental items	
	OL		Recall		Recognition	
	M	*SD*	*M*	*SD*	*M*	*SD*
Experimental	9.85	1.50	4.25	2.97	2.60	1.63
Control	10.15	1.53	9.20	1.94	4.55	1.31

A further test was made to determine whether each of the two groups showed a significant loss in retention for the intentionally learned items of List A. Table I indicates that both groups showed a retention loss for List A, but that the difference for the Experimental Group is greater. For the Experimental Group the t for correlated measures was 12.29 ($p < 0.001$) and for the Control Group, 3.56 ($p < 0.001$).

[5] A. L. Edwards, *Statistical Methods for the Behavioral Sciences*, 1954, 420.

It will be noted that the present design does not provide *OL*-scores for the incidental items. Only scores of retention are possible. Table I reveals that the Control Group surpassed the Experimental Group in the recognition of incidental items. The value of t in this instance was 4.15 ($p <$ 0.001). Retention of the incidental items clearly was significantly greater under the control condition. Thus, *RI* effects in incidental learning have been demonstrated in the present design.

Summary

This study investigated retroactive inhibition for incidentally learned materials. An experimental design was used which differed from that used in previous studies. Two groups of *S*s were given three trials of intentional learning on a list of names of common objects. Each word was accompanied by a different two-digit number (incidental items). Following *OL*, one group received a three-trial presentation of another series of names of common objects, while the other worked on a simple puzzle for the length of time required for *IL*. The interpolated words were not accompanied by incidental items. Subsequently, both groups were tested for retention of the intentionally learned words and the incidental numbers. The results indicate that the interfering effects of *IL* extend to both intentionally and incidentally learned materials, even when the interpolated materials differ from the incidental items, and in addition, contain no incidental materials.

APPARATUS

APPARATUS FOR CONDITIONING AND MEASURING THE ACTIVITY OF RATS

By FRED L. ROYER, CLINTON C. BROWN, and WILLIAM LOVE,
VA Hospital, Perry Point, Maryland

The activity wheel has for many years been used extensively as a research instrument to investigate a wide variety of problems in nutrition, drive-level, and reactivity to stress. To meet the special needs of a study undertaken by the senior author, a modification was made of the activity wheel which provides both a measure of total activity and of rate of

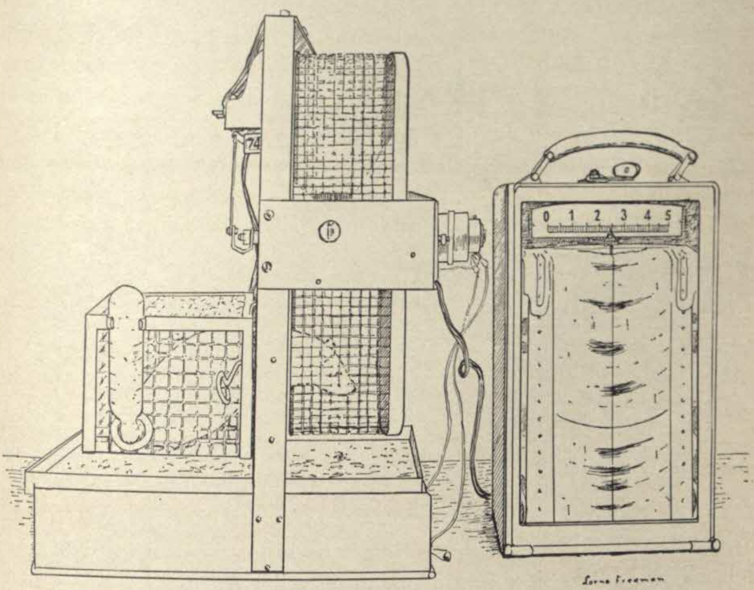


FIG. 1. A DRAWING OF THE MODIFIED INSTRUMENT AND THE RECORDING MILLIAMMETER.

running. The resultant instrument greatly increases the amount of information about activity over standard counts made from the mechanically activated counter summed over a period of time.

The modified instrument is shown in Fig. 1. The Wahmann Activity

Cage is the basic unit. Its axle was geared in a 1:2 ratio to a linear DC tachometer-generator (Model J-36, Eastern Air Devices, Inc., Brooklyn, N.Y.) which generates exactly 2 VDC/100 RPM. This output current is

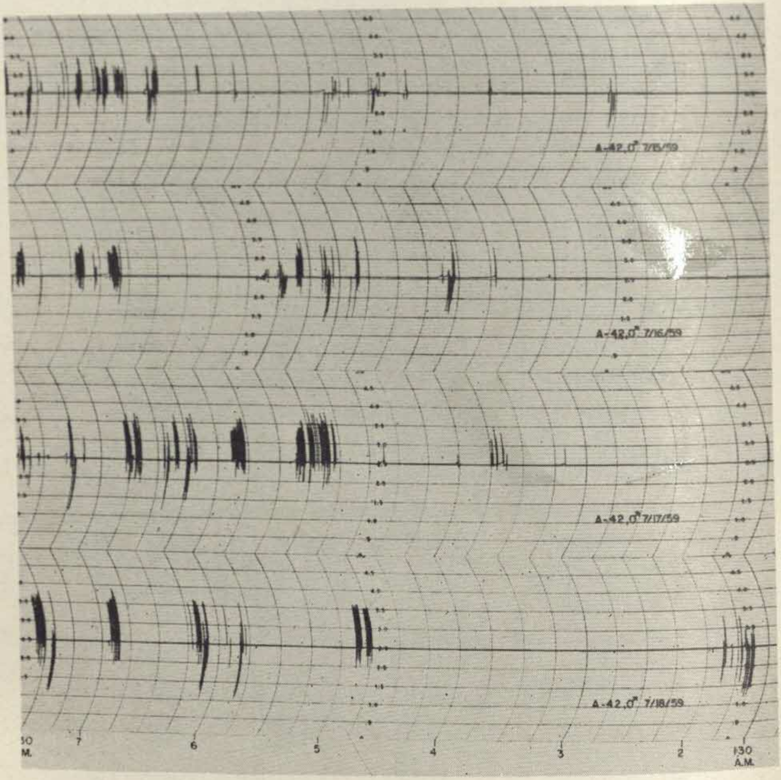


FIG. 2. EXCERPTS OF RECORDS OF A MALE RAT OVER A 4-DAY PERIOD COVERING THE PERIOD OF 1:30 TO 7:30 A.M. DAILY. THE PERIODICITY OF THE ANIMAL'S MAJOR ACTIVITY MAY BE OBSERVED.

divided by a 2.5K ohm series potentiometer which acts as a sensitivity control and the output is recorded on an Esterline-Angus 2½-0-2½ MA Recorder without amplification.

To calibrate, a VTVM is placed across the output of the generator and the wheel is spun until the desired voltage is indicated. The sensitivity control is then adjusted to produce the desired deflection of the recording milliammeter. On the records reproduced in Fig. 2, 3, 4, 5, each division equals 5 RPM.

Typical information yielded by the apparatus. Since the tachometer-

generator produces a DC voltage, the direction of its spin will give either a positive or negative voltage and preferences of the rat for running in one direction or another can be determined. This is shown clearly in Fig. 2, 3, and 4.

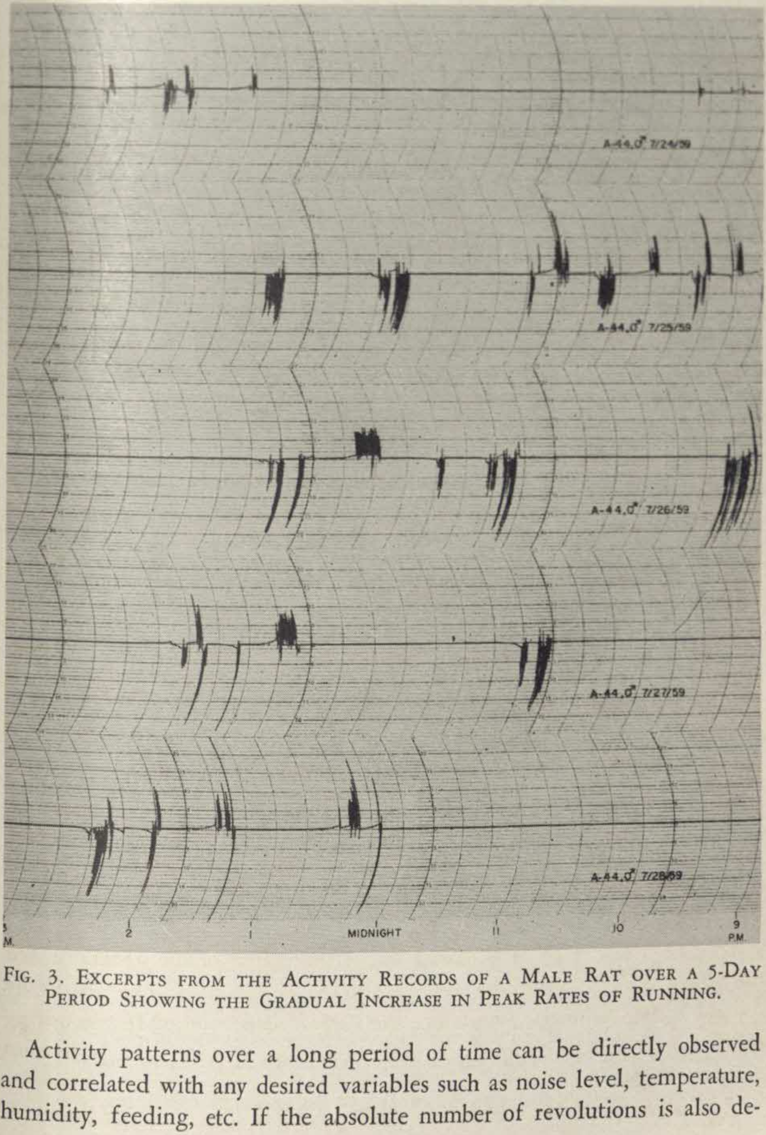


FIG. 3. EXCERPTS FROM THE ACTIVITY RECORDS OF A MALE RAT OVER A 5-DAY PERIOD SHOWING THE GRADUAL INCREASE IN PEAK RATES OF RUNNING.

Activity patterns over a long period of time can be directly observed and correlated with any desired variables such as noise level, temperature, humidity, feeding, etc. If the absolute number of revolutions is also de-

sired on the record, a simple counter can be devised from a single-point cam mounted on the axle so as to actuate a microswitch. This actuates a stepping relay wired to an event marker so that after the desired number of revolutions and steps the marker will record a "pip" on the record.

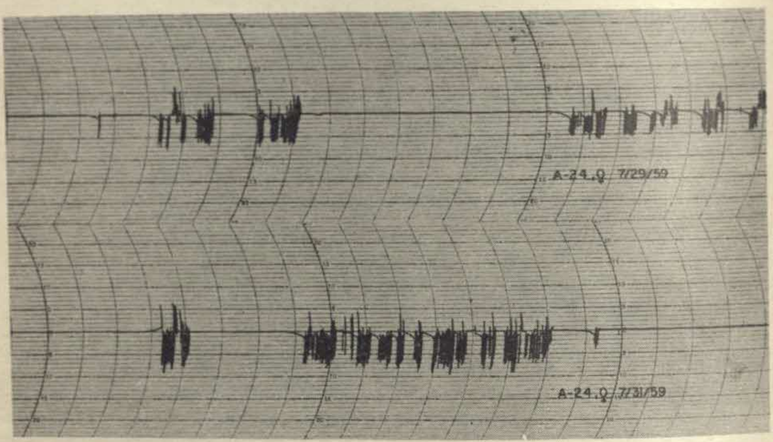


FIG. 4. EXCERPTS FROM RECORDS OF ACTIVITY OF A FEMALE RAT.
The low-rate activity over long periods of time is in contrast to the more frequent high-rate for short time periods of other rats.

The periodicity of spontaneous activity may be determined from the records. In Fig. 2 the activity of one rate from 1:30 to 7:30 A.M. for four consecutive days is presented. The animal had food *ad lib.* and was in a small sound-shielded enclosure. The lights were on for 24 hours a day on July 15, 16, and 18; they were off for 24 hours on July 17.

Changes in the pattern of running may be observed by this apparatus.

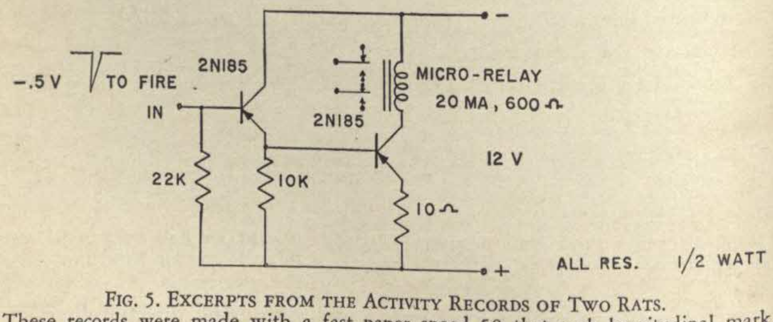


FIG. 5. EXCERPTS FROM THE ACTIVITY RECORDS OF TWO RATS.
These records were made with a fast paper speed 50 that each longitudinal mark represents 15 seconds. Note the oscillation, marked by an arrow, which is produced by the animal "rocking" the wheel.

A gradual increase in peak-rates of running may be observed in Fig. 3. The series of records shown were selected from a five day period under conditions similar to those mentioned above, except that the lights were on continuously. The excerpt from the record of July 24 indicates a maximal peak-rate of 40 RPM; of July 25, 85 RPM; and on July 26 the peak-rate reaches 125 RPM. The high rate was maintained.

Fig. 4 emphasizes individual differences in running rates. The two samples from a three day period show a consistent long-time, slow-rate

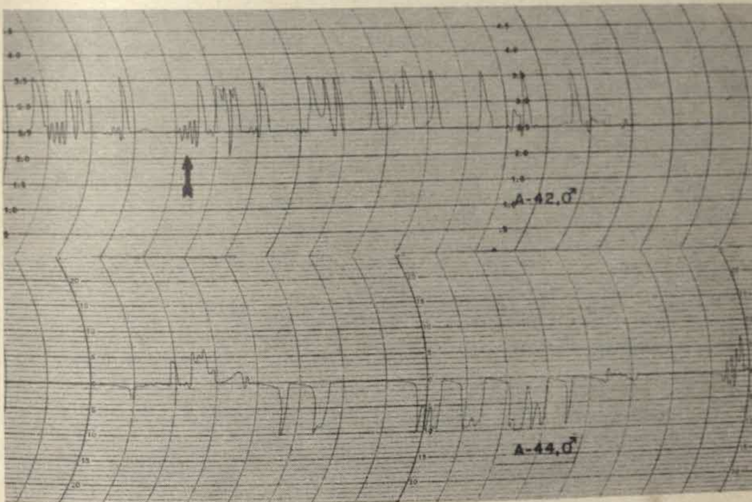


FIG. 6. CIRCUIT DIAGRAM OF THE ELECTRONIC SWITCH USED FOR KEYING A RELAY WHEN THE RATE OF RUNNING EQUALS OR EXCEEDS A SPECIFIED RATE.

activity pattern which differs from the usual pattern of short bursts of high-rate activity.

Additional information may be obtained about running behavior during short time intervals by increasing the chart speed. In Fig. 5 the paper speed is sixty times as fast as in the previous records, *i.e.*, every time division on these represents 15 sec. instead of 15 min. The sinusoidal waves crossing the zero-line, marked with an arrow, are produced by rocking; *i.e.*, the animal comes to an abrupt stop and rocks in the wheel.

Modification for conditioning procedures. An electronic switch, Fig. 6, can be added to the output of the tachometer-generator. This switch keys a relay which controls the presentation of a stimulus. Thus, when the animal's running equals or exceeds a specified rate, the switch will

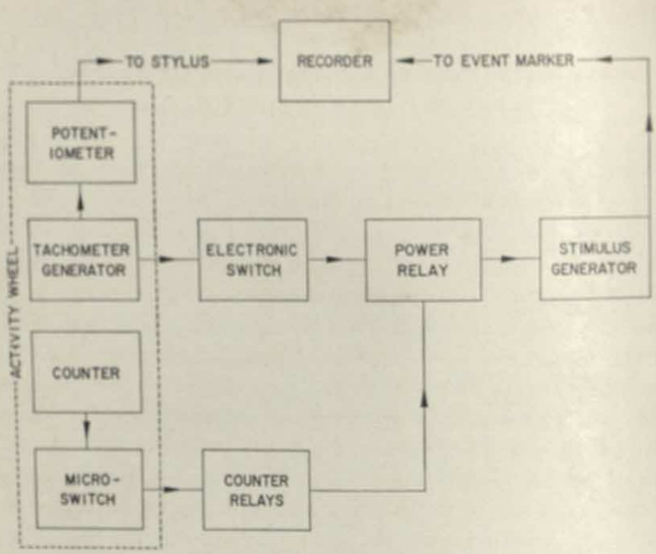


FIG. 7. BLOCK DIAGRAM OF THE CIRCUIT.

either initiate or discontinue a stimulus. When appropriate relays, counters, and timers are added to the above, various reinforcements can be made contingent upon the animal's running behavior. See the block diagram in Fig. 7.

CONTINUOUS DIRECT MEASUREMENT OF APPARENT SKIN CONDUCTANCE

By DAVID T. LYKKEN, University of Minnesota, and NORMAN ROTH, Medtronic Company, Minneapolis, Minnesota

The choice of a unit in which to express the galvanic skin response (*GSR*) or the slower 'tidal' changes of the static skin-resistance has been a matter of considerable debate. Since the psychologist is interested in electrodermal changes merely as indicants of some (as yet unspecifiable) central excitatory process, the problem is to find a unit which will vary linearly with the amplitude of this central event. It now seems clear that the volar (palmar) *GSR* results from bioelectric changes attendant upon the excitation of the eccritic sweat glands.[1] Darrow has shown that the apparent electrical *conductance* of the skin is a linear function of the rate of secretion of the sweat glands.[2] Except under conditions of extreme sweating, however, the rate of secretion in a given area is a function of the proportion of glands in that area which are active at the time.[3] A reasonable inference as to the sequence of events might be as follows: a central excitatory process produces increased excitation in the region of the electrode thus activating a larger proportion of sweat glands. Every gland activated serves as an additional path of low resistance through the high resistance, cornified layer and *stratum lucidum;* hence, each newly activated gland will add an approximately equal increment to the apparent electrical conductance of the skin in that region.[4] Thus, it would seem that apparent skin conductance, rather than resistance, should be more nearly linearly related to the magnitude of the originating central process. Many recent investigators using the *GSR* as a dependent variable have employed change in apparent conductance as their unit of measure. The technique commonly used for electrodermal measurement, however, is to pass a regulated con-

[1] Contraction of the myoepithelial lining of the secretary tubules may also be implicared in the *GSR*. U. Ebbecke, Arbeitsweise der Schweissdrüsen und sudomotorische Reflexe bei unmittelbarer Beobachtung mit Lupenvergrösserung, *Arch. ges. Physiol.*, 253, 1951, 333-350.

[2] C. W. Darrow, The significance of the *GSR* in the light of its relation to quantitative measures of perspiration, *Psychol. Bull.*, 31, 1934, 697-698.

[3] W. C. Randall, Sweat gland activity and changing patterns of sweat secretion on the skin surface, *Amer. J. Physiol.*, 147, 1946, 391-398.

[4] That the activated gland does provide a low resistance path has been well established although the mechanism is less clear; *cf.* Y. Kuno, *Human Perspiration*, 1956, 12.

stant current through the skin and then to record the variations of the voltage drop across the skin; by Ohm's Law, these voltage fluctuations give a linear measure of changes in apparent skin resistance (*SR*). To convert such records into conductance units, one must divide each change in resistance (*GSR*) by the product of the *SR* at the instant of the start of that *GSR* times the *SR* at the peak of that *GSR* deflection. This is a laborious procedure at best. At worst, it introduces considerable error of measurement, since the error in conductance units will tend to exceed the error of measurement of *SR* or of *GSR*, whichever is larger. With typical electrodes,[5] the resistance *GSR* may average about 1% of the *SR*, so that a (good) measurement accuracey of 1% in *SR* may yield 100% error in resistance *GSR* as measured from the same record and, hence, extremely large errors in the calculated values of change in conductance. Recording of *SR* and *GSR* on two separate channels is a possible but expensive solution. Automatic range-changing devices are another alternative but are notoriously prone to go into action right in the middle of a response, causing a loss of valuable data. In either case, the labor of converting to conductance units is still formidable.

A more promising instrumental approach is to apply, instead of a constant current, a constant external voltage to the skin and then to measure the changes in current which, again by Ohm's Law, will now provide a linear record of skin conductance directly. (It is true that Ohm's Law does not necessarily apply to the skin, which behaves like a normal conductor in parallel with an electrolytic cell. It can be shown, however, that apparent conductance does in fact vary directly with current for constant applied voltages in the range which would be used for measurement, *i.e,* that Ohm's Law *does* apply at least within this range.)[6] This principle has been employed several times in the past for measuring static skin conductance by the simple expedient of attaching a battery and low resistance galvanometer in series with *S*. Unfortunately, this simple apparatus is not easily adapted to the making of permanent recordings of the faster and shallower conductance *GSR*'s. We have therefore designed a more complex but more versatile instrument for this purpose.

Apparatus. With the device described here, a constant potential of 10, 1.0, 0.1, or 0.01 v. is impressed across *S*, that range being chosen which, for the given *S*, will pass a current of less than 10 μamp. A transistor amplifier which has an input impedance of about 100 Ω is in series with *S* and amplifies this varying current enough to drive a panel meter and an external recording milliameter. Because current

[5] D. T. Lykken, Properties of electrodes used in electrodermal measurement, *J. comp. physiol. Psychol.,* 52, 1959, 629-634.
[6] Lykken, *op. cit.,* 632.

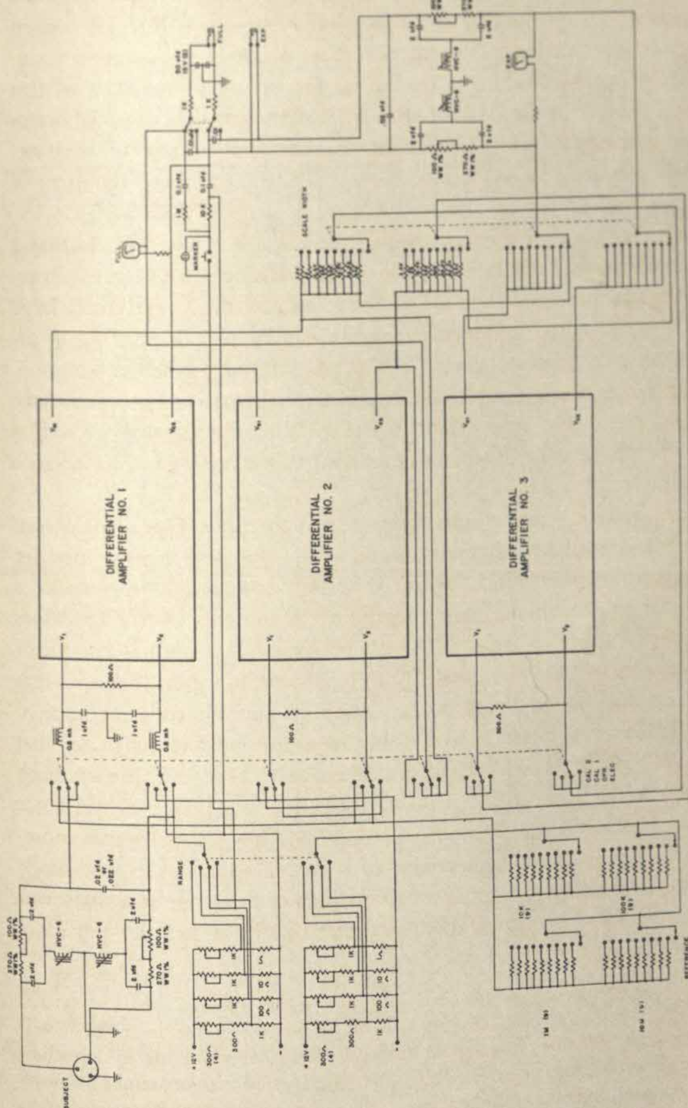


FIG. 1. AN INSTRUMENT FOR DIRECT MEASUREMENT OF THE APPARENT ELECTRICAL CONDUCTANCE OF THE SKIN
(See text for explanation)

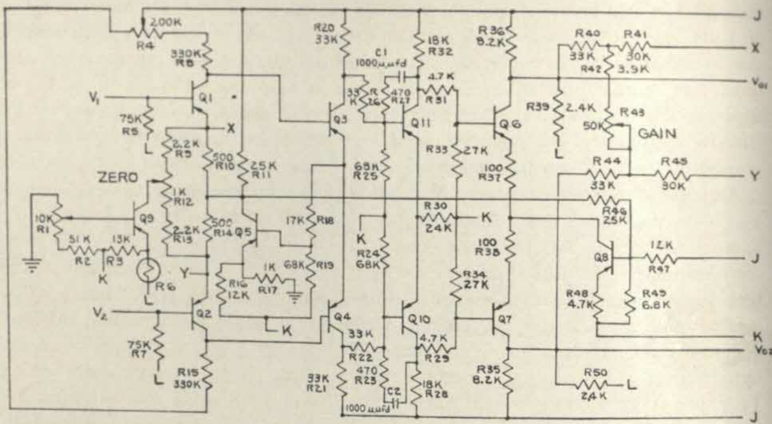


FIG. 2. DETAILED CIRCUIT OF DIFFERENTIAL AMPLIFIER SECTION
(Three of these required)
All transistors are 2n359; J = −30 v.; K = +30 v.; all resistances are in ohms; and all resistors are within 1% tolerance.

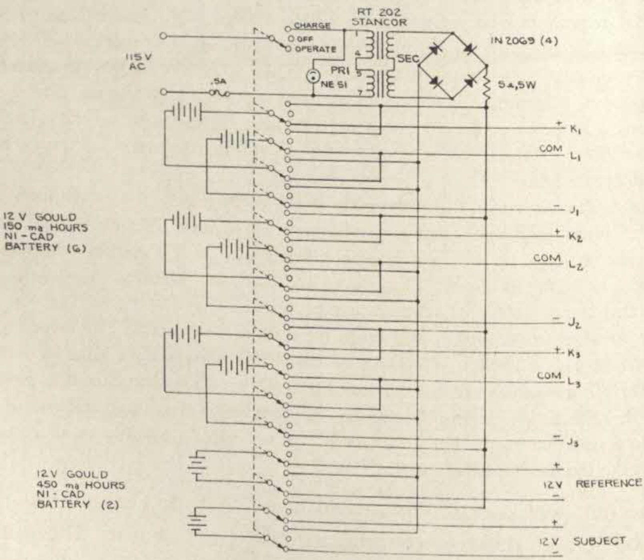


FIG. 3. POWER SUPPLY
(Nickel-cadmium storage batteries and a trickle-charging circuit)

through S is held below 10 μamp., electrode polarization effects are minimized. Use of such low current-levels with constant-current circuits is normally objectionable due to the artifacts which may be introduced by changes in skin potential; *e.g.* at 5 μamp., a potential *GSR* of 5 mv. will appear as a spurious change of 1000 Ω of apparent skin-resistance. The extremely low input impedance of the present circuit effectively eliminates the endogenous skin potentials by shunting them, making it possible to use low current levels without producing artifacts of this kind.

An identical reference amplifier has at its input a reference conductance or 'dummy S' whose conductance is adjustable from zero to 1000 μomhos in increments of 0.1 μomhos. By setting this reference conductance to a value about equal to S's static conductance, the residual difference between the outputs of the two amplifiers can then be fed to a third amplifier which then expands the small *GSR* fluctuations by a factor of 1, 2, 5, 10, 20, 50, or 100. The output of this scale-expanding amplifier operates a second panel meter and can also be fed to a second channel of an external recording milliameter. Thus, for example, using the 0-10 μomho range setting (1 v. across S), a resistance of 200,000 Ω will give a half-scale reading on the *SC* meter and a change from this resistance of 4000 Ω (0.1 micromhos) will give full-scale deflection of the *GSR* meter at maximal (x 100) expansion.

The instrument was designed to operate with a Sanborn #150 DC amplifier and recording milliameter and gives an output for full-scale meter deflection of 0.5 v. on both channels. It can be used either with a two-channel recorder or, by interposing an electronic switch (*e.g.* the Sanborn 'Triplexer'), both the static *SC* and the *GSR* records can be written on a single channel. Or, an adequate one-channel record can be obtained by using a manual switch on the panel which alternately connects the static *SC* or the *GSR* output to the recorder. In this arrangement, one merely takes a reading of static *SC* as often as needed. Since the *GSR* output is recording conductance directly, it is not necessary to measure the static level at the onset of every *GSR* as one must do when recording resistance for later conversion to conductance units.

Circuit. The circuit of the instrument is shown in Fig. 1. Low-pass filters are provided at both input and output to eliminate 60~ noise. Tuneable high frequency filters are included to act against the interference of *TV* signal which can be a nuisance in some locations. The three amplifiers are identical, each being highly stable and compensated for temperature variations. They have a gain of 500 and the ability to attenuate in-phase *AC* noise by a factor of 50,000. The amplifier circuit is given in Fig. 2, where it should be noted that all resistors must be accurate to 1% and all transistors are the standard type 2N359. The instrument is powered by a set of 12-v. nickel-cadmium batteries which give good voltage stability as well as complete isolation from the AC line. This power supply together with a circuit for recharging the batteries is shown in Fig. 3.

The unit will operate for several hours on fully charged batteries and may be left on 'charge' continuously when not in use. The instrument should be turned on at least 15 min. before being used to allow temperatures and battery voltage to reach stable levels. Once calibrated, the completed unit is extremely sensitive, accurate, and convenient to use.

AN APPARATUS FOR SELF-ADJUSTMENT OF THE INTERVAL IN DELAYED ALTERNATION BY MONKEYS

By MURRAY E. JARVIK and STEPHAN CHOROVER,
Albert Einstein College of Medicine

The test of delayed alternation was early proposed as a measure of immediate memory similar to spatial and non-spatial tests of delayed response. It is a task which has been shown to be sensitive to frontal brain lesions as well as to drugs.[1] Even though it is a task which involves memory, it is possible, however, that other factors (such as changes in perception or motivation) may be responsible for the deficits in performance reported. The present technic was developed in connection with a project to study the effects of drugs on behavior in monkeys. It was designed to determine whether retention is more closely related to poor performance during the delay- or the predelay-periods.

The apparatus used for determining the delay-threshold consists of the following. The animal is restrained by the neck in a test chamber. In front of him there is a panel containing two toggle switches, signal lamps, and a nozzle for delivering liquid reward. A monkey, deprived of water for more than 20 hr. is placed in the chamber. It can obtain water from the nozzle by depressing the toggle switches alternately when the 'go' signal, an overhead light, is turned on. Every correct response, that is, a response to the correct switch at the correct time, is rewarded by 0.3 cc. of sweetened water. Control of the timing sequence is arranged by means of a circuit containing cascaded counting relays so that every time the animal makes a correct response the succeeding delay is lengthened by 1 sec. The present arrangement permits delays up to 10 sec., but additional stepping switches will allow multiplication of this interval. Should the animal press during the delay interval, that is, make a premature response while the overhead light is off, the delay timer is reset to zero and the animal must wait for an additional time equal to the delay before the 'go' signal reappears. The length of the delay, however, is not changed by this anticipatory response.

[1] Mortimer Mishkin and K. H. Pribram, Analysis of the effects of frontal lesions in monkeys; I. Variations of delayed alternation, *J. comp. physiol. Psychol.*, 48, 1955, 492-495; M. E. Jarvik and Stephen Chorover, Effects of lysergic acid diethylamide upon certain aspects of memory (delayed alternation) in monkeys, *Fed. Proc.*, 17, 1958, 1.

He can, of course, prolong the delay interval indefinitely in this way, but these early responses tend to fall out relatively quickly in learning.

On the other hand, each time the animal makes a response to the incorrect key, but at the appropriate time, this is followed by an unrewarded delay equal to the maximal interval, and thereafter the stepping switch resets so that the minimal delay is again effective. Should he then respond correctly the delay for the next trial will be 1 sec. longer. In this way the maximal delay reached in any sequence of trials is a function of the probability of the animal making an error by responding to the wrong key on the previous trial. It can be seen that this procedure is similar to the psychophysical method of limits or of serial exploration, except that only an ascending order of difficulty is used.[2]

The responses of a given animal are recorded on a continuously moving

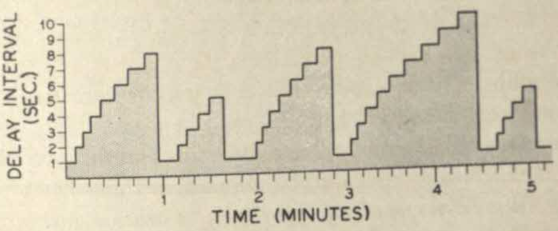


FIG. 1. SAMPLE RECORD OF PERFORMANCE ON DELAYED-ALTERNATION

chart which indicates the level of difficulty at which a given response is made. A 5-min. sample of performance is shown in Fig. 1 where the value on the ordinate indicates the duration of the delay in seconds and time is the abscissa. Each correct response produces an upward deflection of the pen and an accompanying delay which is 1 sec. longer than that preceding it. Every error in the choice of the key causes the timing mechanism to reset and the pen to return to the base line. It can be seen in Fig. 1 that in the particular 5-min. period shown the animal made 36 responses, 5 of which were errors. These errors occurred twice at 5 sec., twice at 8 sec., and once at 10 sec. Such a method of recording is particularly useful in studying the effects of drugs upon this behavior.

Fig. 2 shows the effects of various delays on the accuracy of performance in this test. It can be seen that this animal was able to perform the task at a high level of accuracy for the short intervals but that accuracy

[2] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1954, 133-159.

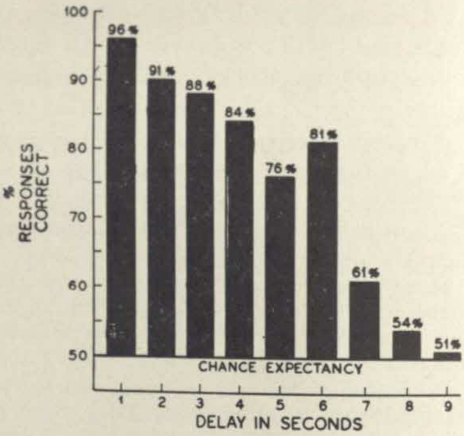


FIG. 2. TYPICAL PERFORMANCE SHOWN AS PERCENTAGE OF RESPONSES CORRECT AFTER VARIOUS DELAYS

diminished with increasing delay until at 9 sec. the percentage of correct responses was very close to that which could be expected by chance alone.

It is possible by this method to have the animal set the level of difficulty of the problem close to his particular threshold. Thus, the threshold can serve as a sensitive measure of physiologically or pharmacologically induced changes, whereas accuracy of performance obtained after a fixed delay may be much less sensitive.

APPARATUS NOTES

EYEBLINK CONDITIONING: PHOTOELECTRIC PICK-UP AND *EEG* RECORDING

Franks and Withers described a photoelectric method of recording movements of the eyelid.[1] Their chief innovation was a photoelectric cell mounted on the rim of spectacles which converted reflected light from the eye and lid into electrical impulses which varied with the dropping of the lid. They fed the output of the photoelectric cell into an amplifier of their own design which drove their recording milliameter. Their method is accurate, relatively inexpensive, and convenient; it does not place undue restrictions upon *S*'s movements nor require that electrodes, paper eyelashes or other foreign objects be attached to the skin.

Since, however, their cell operates off incidental light, the sensitivity varies with room illumination, the range being limited unless special steps are taken, *e.g.* utilizing infra-red light; also, the tracing fluctuates as *S* turns his head unless the illumination is homogeneous. A simple modification of their apparatus described here reduces restrictions as to lighting and clears the visual field. An electroencephalograph *(EEG)* circuit, standard equipment in most laboratories, has furthermore been substituted for the custom-built amplifier-recorder.

In our modification, compressed air is fed through a polythylene tube into an assembly of copper tubes constructed into the shape of a 'T' (Fig. 1). The air enters the body of the 'T' (5 mm. in internal diameter and 70 mm. long) and escapes through one radial arm (1 mm. diameter, 15 mm. long) which is pointed into *S*'s left eye. The other arm of the 'T' (7 mm. diameter, 20 mm. long) houses a miniature, pre-focused flashlight lamp held in place by an air-tight seal of plastic tape.

A small photoelectric cell (International, Self-Generating Selenium, B2M), which is quite inexpensive and only 0.7 × 0.4 × 0.04 in. in size, was chosen to replace Franks and Withers' more bulky tube-type. The cell is flat, wafer-shaped; sensitive on only one side, it does not require shielding.

The air-puff and light assembly are integrated into a unit with the photoelectirc cell by running the 1-mm. tube through a hole drilled in the center of the cell, thereby reducing obstruction in the visual field and assuring the proper alignment of the reflected light. A thin layer of plastic insulates the copper from the cell. Wires lead off from the lamp to a 1.5-v. battery and from the cell to the input of one *EEG* channel.

The unit is mounted onto the rim of the spectacles with a small bolt. An expansible washer exerts enough pressure to maintain the unit in proper position but allows enough play to adjust the beam of light and air-puff to strike the most desirable point of *S*'s eye. A rubber band passing behind *S*'s head holds the frame firmly in place. Both lenses are removed from the spectacle frames.

[1] C. R. Franks, and W. C. R. Withers, Photoelectric recording of eyelid movements, this JOURNAL, 68, 1955, 467-471.

The light from the miniature bulb forms a circle on the sclera of the eye. Brighter than the incidental reflected light, it makes the functioning of the cell largely independent of room-illumination; light and dark areas have little effect and the average illumination-level may be very-bright to nearly-dark. Little light enters the pupil; under normal conditions *S* may not realize that the bulb is on. Lamp black daubed on the eyelid would increase the sensitivity of the apparatus.

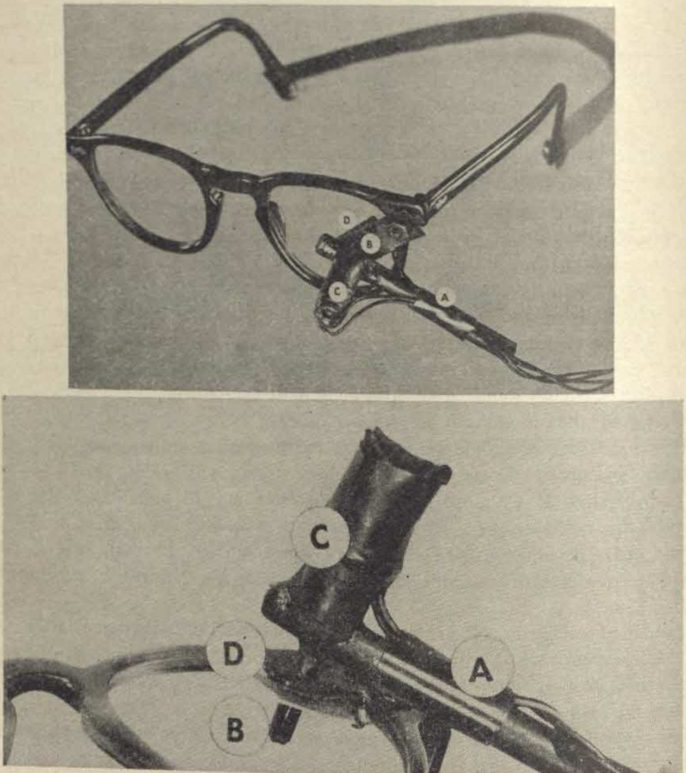


FIG. 1. SPECTACLE FRAMES WITH CONDITIONING UNIT MOUNTED ON RIM
Air is forced through (A), escapes through (B). (C) contains lamp which shines through (B); reflected light, varying with position of lid, induces voltage in photoelectric cell (D). Wires lead from battery to lamp and from photoelectric cell to the input of the electroencephalograph.

The substitution of the *EEG* for amplifying and recording has the advantages of sparing the laboratory the expense of constructing specialized equipment. It is well-calibrated as to amplification and the rate of paper feed. Its multiple channels permit the simultaneous recording of other variables. Of perhaps the most practical importance, technicians are available who are thoroughly familiar with the equipment and who know the idiosyncrasies of the particular machine.

A DC amplifier, were one available, would be preferable to the *EEG*. After the blink is completed and the lid has raised, the pen returns to the neutral position instead of remaining at its peak—the graphic tracing does not correspond in this respect to the lid-position. If the lid were held closed for a short time, the same drift back to mid-line would occur. Franks and Withers' apparatus suffers from the same limitation. The actual movement of the eyelid is however, faithfully recorded.

V.A. Hospital, Lexington, Kentucky
University of Kentucky

RICHARD M. GRIFFITH
CHARLES H. SPARKS

TEMPERATURE-CONTROL FOR A GROUP OF SMALL AQUARIUMS

Control of temperature is a problem when a large number of tropical fish must be kept in individual tanks. To have an immersion heater for each tank is costly and cumbersome. To heat the air in the room is feasible only if *E* is prepared to tolerate a certain amount of discomfort and if all *S*s in a given room require the same temperature. The method to be described here has none of these limitations. It is convenient and inexpensive; it involves no discomfort to *E*; it permits different groups of tanks in the same room to be kept at different temperatures.

The tanks are set on wooden shelves. To each shelf is stapled a layer of soft asbestos paper and then a layer of heavy kitchen-grade aluminum foil. Vinyl-jacketed heating cable (G.E. HS 100) is run back and forth on top of the aluminum foil in the longer direction of the shelf, with a spacing of about 1 in. between strands. The power consumption of this cable is 5-w./ft. Insulated wire staples are used to keep the strands in place. The tanks rest directly on the heating cable and thus are heated from below. A thermostat, in the circuit with the heating cable, and its plastic covered sensing element, immersed in one of the tanks in the group, serve to regulate the temperature of all tanks in the group.

Depending on the minimal room temperatures anticipated, the distance between strands of cable may be increased or decreased. The exact arrangement to meet any given set of requirements might best be determined empirically. In our own application, 200 ft. of cable serve to maintain 30 2-gal. tanks at 82°±2°F. in a room in which air temperatures have fallen below 40°F. for periods as long as 12 hr. The margin of safety in this arrangement is indicated by the on-off ratio of the heating cable, which is on the order of 20% at an air temperature of 70°F.

The New School for Social Research

ROBERT M. STEINMAN
JEROME WODINSKY

NOTES AND DISCUSSIONS

SOME OBSERVATIONS REGARDING THE EXPERIENCES AND BEHAVIOR OF THE BAMBUTI PYGMIES

The identity of the BaMbuti Pygmies of the Ituri Forest in the Congo with the forest itself goes beyond their social life; they are also psychologically conditioned by their environment. This can best be illustrated by some observations that I made during a recent field trip in their country.

Distance- and size-perception. At the end of a particularly long and tiring period of trekking through the forest from one hunting group to another, I found myself on the eastern edge, on a high hill which had been cleared of trees by a missionary station. There was a distant view over the last few miles of forest to the Ruwenzori Mountains: in the middle of the Ituri Forest such views are seldom if ever encountered. With me was a Pygmy youth, named Kenge, who always accompanied me and served, amongst other capacities, as a valid introduction to BaMbuti groups where I was not known. Kenge was then about 22 yr. old, and had never before seen a view such as this. He asked me what the "things" before us were (referring to the mountains). "Were they hills? Were they clouds? Just what were they?" I said that they were hills bigger than any in his forest, and that if he liked we would leave the forest and go and see them and have a rest there. He was not too sure about this, but the BaMbuti are an incorrigibly curious people and he finally agreed. We drove by automobile in a violent thunderstorm which did not clear until we entered the Ishango National Park at the foot of the mountains and on the edge of Lake Edward. Up to that moment from the time we had left the edge of the forest, near Beni, visibility had been about 100 yd.

As we drove through the park the rain stopped and the sky cleared, and that rare moment came when the Ruwenzori Mountains were completely free of cloud and stood up in the late afternoon sky, their snow-capped peaks shining in the sun. I stopped the car and Kenge very unwillingly got out. His first remark was to reiterate, what he had been saying ever since the rain stopped and we could see around us, that this was a very bad country, there were no trees. Then he looked up at the mountains and was completely unable to express any ideas—quite possibly because his language had no suitable terms, being limited to the experience of a strictly

forest people. The snow fascinated him, he thought it must be some kind of rock. More important, however, was the next observation.

As we turned to get back in the car, Kenge looked over the plains and down to where a herd of about a hundred buffalo were grazing some miles away. He asked me what kind of insects they were, and I told him they were buffalo, twice as big as the forest buffalo known to him. He laughed loudly and told me not to tell such stupid stories, and asked me again what kind of insects they were. He then talked to himself, for want of more intelligent company, and tried to liken the buffalo to the various beetles and ants with which he was familiar.

He was still doing this when we got into the car and drove down to which the animals were grazing. He watched them getting larger and larger, and though he was as courageous as any Pygmy, he moved over and sat close to me and muttered that it was witchcraft. (Witchcraft, incidentally, is known to the BaMbuti only through association with the Bantu. They have no similar concept of the supernormal). Finally when he realized that they were real buffalo he was no longer afraid, but what puzzled him still was why they had been so small, and whether they *really* had been small and had suddenly grown larger, or whether it had been some kind of trickery.

As we came over the crest of the last low hill, Lake Edward stretched out into the distance beyond, losing itself in a hazy horizon. Kenge had never seen any expanse of water wider than the Ituri river, a few hundred yards across. This was another new experience difficult for him to comprehend. He again had the same difficulty of believing that a fishing boat a couple of miles out contained several human beings. "But it's just a piece of wood," he protested. I reminded him of the buffalo, and he nodded unbelievingly.

Later we went all over the National Park with one of the African guides. He and Kenge conversed in KiNgwana, the *lingua franca* of the area, and Kenge was constantly looking out for animals and trying to guess at what they were. He was no longer afraid or unbelieving; he was trying to adapt himself, and succeeding, to a totally new environment and new experience.

The next day he asked to be taken back to the forest. He reverted to his original argument. "This is bad country, there are no trees."

The inability of the BaMbuti to correlate size-constancy and distance had never even struck me as a possibility. In the forest, vision is strictly limited to a matter of yards, the greatest distance one can see, when up a tree looking down onto a camp, being a hundred feet or more below. Kenge

was, however, a sophisticated and well travelled Pygmy. He had been with me a long time, had travelled along roads where he could see for as much as a quarter of a mile, and had seen aircraft and knew that they contained people. Such instances, however, were rare, and on the whole his experience of visual distance was limited to the relatively slight diminution of size in seeing a person or people walking along a road a quarter of a mile away. He had seldom seen any animal from further away than a few yards, he had never seen any boat bigger than a dug-out canoe, and that no further away than a few hundred feet.

Number-perception. Size-perception is, however, only one of many phenomena of interest to the psychologist. The Pygmy, unless he is one who has constant dealings with the Bantu, is unable to count above four. He has, however, such an eye for patterns that, for example, if several arrows are taken from a bunch, he can detect the reduction and can usually replace the correct number withdrawn to bring the bunch to its original size. In a gambling game (*panda*) common in the region, up to 40 or so pebbles, seeds, or beans are thrown onto a mat. In a single glance the Pygmy can tell you if they form a multiple of four, or how many—one, two, or three—have to be added to make it into such a multiple. The game is a test of skill in number perception and manipulation. Spare beans are concealed between the fingers and toes, and as a player makes his throw, while the beans are still rolling on the mat, he has already made his calculation and added the requisite number from his concealed reserve to bring the total to the winning multiple.

Art: (1) Visual. Another phenomenon worthy of study, and again associated with environmental influence, is the almost total lack of any form of physical art. The BaMbuti refer to white, black and red by color names, for other colors they make comparisons—"like leaves," "like leopards," instead of "green" or "yellow." They use red or blue-black dyes in the crude decoration of their bark cloths, smearing the dye on with their fingers. More complicated are the designs painted on the bodies of the girls and women, using the black stain obtained from the gardenia fruit. Except for these decorations, visual art is lacking. Wooden implements are never carved or decorated or even polished. Perhaps the word of the BaMbuti is too close around him, too confined and colorless, too much lacking in variety, to produce a visual art.

(2) Auditory. In contrast to this lack the Pygmy has the most complex music in the whole of Africa. It is complex not only in terms of rhythm, melody, and harmony (the latter surprising enough in Africa), but also in terms of technique. The BaMbuti can improvise a 15 part liturgy or

canon, with melodies frequently running in parallel seconds, and hold it without the slightest difficulty. When this gets too tame, they divide the melodic line up, note by note, among the performers, each of whom will hoot his note at the appropriate moment. The melody then travels counter-clockwise around the group who may be sitting about a central fire or even in the natural circle formed by their huts, each at his own hearth. There is obvious material here for anyone interested in esthetics, as well as for those who might be more interested in the relatively small part that vision plays in the life of these forest nomads. (Even when hunting, a great deal is done by hearing rather than seeing, and perhaps even smell is more important as a sense. Vision is used by the hunters in the examination of tracks, but the firing of the arrow is often done by sound rather than sight.) I should mention again that music permeates their whole life.

Historical records. The earliest historical records of the BaMbuti, found in a tomb dating from the sixth dynasty in Egypt, places this tribe where it is today, refers to it as forest dwellers, and indicates that song and dance played a great part in the life of its people then just as it does today, over four thousand years later. In the forest there are few forces that stimulate change, and it is probable that the BaMbuti remained for most of this time living much the same kind of life. As recently as three or four hundred years ago, however, the great Bantu migrations forced certain Bantu and Sudanic tribes into the forest. For a number of highly significant reasons the resultant contact has had relatively little effect on the life of the BaMbuti, who consciously and forcefully reject the values of the plains and savannah, and unite in common opposition to the village world of the invaders.

It is a pity that such an exceptional opportunity for the study of a truly primitive people should be missed. In a few years the opportunity will be gone. There is little literature of scientific value available on these people. The references that I give here are those of the greatest interest, but even so are for the most part of general rather than specific value.[1] The work

[1] Martin Gusinde, Die Kongo Pygmäen in Geschichte and Gegenwart, *Acta Nova Leopoldina*, 76, 1942; *Urwaldmenschen am Ituri*, 1948; *Die Twiden. Pygmäen und Pygmoida im Tropischen Afrika*, 1956, P. E. Joset, Buda Efeba, *Zaïre*, 1, 1948, 137-157; Paul Schebesta, *Among Congo Pygmies*, 1933; *My Pygmy and Negro Hosts*, 1936; *Revisiting by Pygmy Hosts*, 1937; *Les Pygmé es du Congo Belge*, 1952; George Schweinfurth, *The Heart of Africa*, 1874; C. T. Turnbull, Initiation among the BaMbuti Pygmies of the Central Ituri, *J. roy. Anthropol. Instit.*, 87, 1957, 191-216; Legends of the BaMbuti, *ibid.*, 89, 1959, 45-60; Some recent developments in the sociology of the BaMbuti Pygmies, *Trans. N.Y. Acad. Sci.*, Ser. II, 22, 1960, 275-284; The *elima*. A pre-marital festival among the BaMbuti Pygmies, *Zaïre*, 2-3, 1960, 175-192; Field work among the BaMbuti Pygmies, *Man*, 60, 1960; *The Forest People*, 1961.

of Schebesta was undertaken a number of years ago, and his later findings, as well as my own, indicate that particularly with regard to his analysis of the Pygmy-Negro relationship he was observing more from the point of view of the village than of the forest. This was due to the fact that it was impossible for him at that time to have access to the BaMbuti except through the offices of the local Negro chiefs. The presence of the Negroes changed the situation, even when in the forest, from a truly forest to a village environment, and the Pygmies reacted accordingly.

True hunters, particularly those who are as heavily conditioned by their environment as are the BaMbuti, are rare. The fact that they are surrounded by so many different cultures, yet have managed to maintain their own cultural integrity is an indication of the depth and vitality of their way of life and thought, however simple and static it may seem on the surface. If they lose their integrity in the next few years, it will not be because of any process of acculturation, but because the forest is no longer theirs and it will have been physically impossible for them to maintain their forest way of life. They are aware of the future that faces them, and while some say "We shall just have to live like the *savages* and plant bananas," the majority say, "When the forest is no more, we shall die." I am afraid it will be the latter.

American Museum of Natural History COLIN M. TURNBULL

DISTANCE-CONSTANCY

It has often been found that with unrestricted viewing conditions two objects which are equal in physical size will be judged equal in size despite differences in their distances from *O*.[1] There is general agreement about this concept of size-constancy. Some confusion, however, has arisen about what is meant by distance-constancy.[2] Both Purdy and Gibson[3] and Smith[4] have demonstrated that a specified space between two points will be judged to remain constant in extent with changes in the distance of that space from *O*. The present aim is to put forward an integrative concept of distance-constancy—in the sense of finding what Boring calls invariant relation-

[1] J. J. Gibson, *The Perception of the Visual World*, 1950.

[2] See S. H. Bartley and H. J. Adair, Comparison of phenomenal distance in photographs of various sizes, *J. Psychol.*, 47, 1959, 289-297. These writers state that "distance-constancy has received very little attention. Accordingly we do not know so much about it, or what might be meant by the term, distance-constancy, itself."

[3] Jean Purdy and E. J. Gibson, Distance judgment by the method of fractionation, *J. exp. Psychol.*, 50, 1955, 374-380.

[4] O. W. Smith, Distance constancy, *ibid.*, 55, 1958, 388-389.

ships.[5] Whereas two objects of the same physical size are judged equal in size despite differences in their distances from *O* (size-constancy), two objects at the same distance from *O* would be expected to be judged equal in distance despite differences in their physical sizes (distance-constancy). In these terms both Purdy and Gibson's and Smith's results can be considered to demonstrate size- or length-constancy.

In the present experiment *O*s judged the relative distances of a pair of objects whose size-combinations were varied factorially. It was predicted that judgments of distance-equality would be independent of the differences in the physical sizes of the two objects.

Testing was carried out in a well-illuminated room 35 ft. long. Few restrictions were placed on *O*'s viewing. *O* was presented with two white circular disks and judged which was the closer to him. The disks were suspended from thin cords separated by 25° at the eye and were 18 in. above eye-level. Three pairs of disks,

TABLE I

MEAN DISTANCES OF COMPARISON-STIMULUS TO WHICH JUDGMENTS OF DISTANCE-EQUALITY ARE GIVEN

Diameter of comparison-stimulus (in.)	Distance and diameter of standard stimulus								
	15 ft.			20 ft.			25 ft.		
	4 in.	8 in.	12 in.	4 in.	8 in.	12 in.	4 in.	8 in.	12 in.
4	14.84	15.16	15.00	20.16	20.00	20.47	25.31	24.84	25.16
8	14.69	15.16	15.47	20.00	19.84	20.31	24.69	25.00	25.31
12	14.56	14.69	14.56	19.84	19.84	20.16	25.00	25.00	24.84

with diameters of 4, 8, and 12 in., and three standard stimulus-distances (15 ft., 20 ft., and 25 ft.) were used.

Each of the 8 *O*s made relative distance-judgments of the 27 factorial combinations of standard stimulus-size, standard stimulus-distance, and comparison stimulus-size. If the comparison-object was smaller than the standard object, the former was initially placed 5 ft. closer to *O* than the latter and moved in steps of 2 ft. 6 in. in the direction *O* indicated until *O* either reversed his judgment or gave a judgment of equality. When the comparison-object was the larger, it was initially placed 5 ft. further from *O* than the standard and again moved in the direction that *O* indicated. In compiling the results, a judgment that the comparison-stimulus at 15 ft. from *O* was closer, but at 17 ft. 6 in. was farther away, than the standard stimulus was treated as a judgment of distance-equality at 16 ft. 3 in.

The distances of the comparison-stimulus at which judgments of distance-equality were given are set out in Table I for the various combinations of the independent variables. The distances of the comparison-stimulus closely approximate those of the standard-stimulus and not the distances of the comparison-stimulus at which the two objects subtend the same vis-

[5] E. G. Boring, Visual perception as invariance, *Psychol. Rev.*, 59, 1952, 141-148.

ual angle. Thus, under unrestricted viewing conditions two stimuli at the same distance from O will be judged equal in distance irrespective of differences in size. If size-constancy can be expressed as:

$$R_{r_s} = f\,(s_1 = s_2)$$

(where R_{r_s} is a judgment of size-equality, s_1 the size of the comparison-stimulus, and s_2 is the size of the standard stimulus), then distance-constancy can be expressed as:

$$R_{r_d} = f\,(d_1 = d_2)$$

(where R_{r_d} is a judgment of distance-equality, d_1 is the distance of the comparison-stimulus from O, and d_2 is the distance of the standard stimulus).

University of Sydney, Australia RAY OVER

'GOALS' AND 'VALUES' REVALUATED

A recent paper by Wiener and Ehrlich presents some data purporting to demonstrate the empirical equivalence of the concepts of 'goals' and 'values.' The data are semantic differential ratings of these two concepts by a sample of 36 therapists and patients. The authors give the average ratings for each scale over these subjects for each of the concepts and assert that "None of the scale-differences was large enough to be statistically significant. The figures graphically demonstrate the similarity with which the two terms are described . . . In view of the similarity . . . we believe that the two concepts, 'values' and 'goals', have similar connative meaning and are used synonymously in practice."[1]

There are two comments to be made on Wiener and Ehrlich's paper and their main assertion. First, the assertion is true only in a very gross sense. Analysis of the differences on each scale between the average rating of 'goals' and 'values' which have been listed in order of size in Table I reveals a systematic and plausible difference in meaning between the two concepts. Secondly, the assertion of no "statistically significant" differences is unsupported by publication of the standard errors of the differences to which the authors refer, and is, on the basis of the data which are reported, in some doubt.

The differences listed in Table I reveal that 'goals' are actually rated by these subjects as more *timely, competitive, new, conventional, intentional, hard, active, selfish* and *voluntary* and less *inner* than 'values'! This set of differences in rated meaning corresponds closely with the dictionary differ-

[1] D. N. Wiener, and Danuta Ehrlich, 'Goals' and 'values,' this JOURNAL, 73, 1960, 617.

ence in meaning between the two concepts. 'Goals' are defined as "objects or ends that one strives to attain" while 'values' are defined as "acts, customs, or institutions regarded in a particular, especially favorable way."[2] The lack of difference between the two concepts on evaluative scales is surprising, but the observed differences fit well with the more immediate, practical and action-oriented nature of 'goals.'

The omission of standard errors is always unfortunate. The importance of a statistic such as an average cannot be adequately determined without information about its precision such as that given by its standard error. Since, in spite of the authors' assertions, the actual differences in average

TABLE I

THE DIFFERENCE IN MEANING BETWEEN 'GOALS' AND 'VALUES'

(Data from Wiener and Ehrlich's Figs. 1 and 2. Scales reversed to make all differences in favor of 'Goals' positive.)

Scale	Mean rating of		Mean Diff.
	'Values'	'Goals'	
timely/untimely	3.2	5.5	2.3
competitive/coöperative	3.3	5.3	2.0
new/old	3.5	4.5	1.0
outer/inner	2.4	3.3	0.9
conventional/unconventional	4.1	4.9	0.8
intentional/unintentional	4.9	5.7	0.8
hard/soft	4.8	5.6	0.8
active/passive	4.6	5.3	0.7
selfish/unselfish	3.5	4.1	0.6
voluntary/compulsory	4.5	5.1	0.6

ratings are quite meaningful, it is particularly unsatisfying not to be informed of the precision of these averages in this paper.

It is possible to estimate what this precision may be by making an assumption about the distribution of ratings per scale. This estimate, as I will now show, casts some doubt on the authors' assertion of no "statistically significant" differences.

If there was any agreement among the 36 raters on the meaning of a concept, then the distribution ratings of this concept on a given scale was heaped up around its average. The variance of a uniform distribution over a seven-point scale is 4 and so we can assume that the sample variance of ratings for any scale in the Wiener and Ehrlich study, on which there were any signs of group agreement, was less than 4.

If the variance of ratings was less than 4, then the standard error of an average of 36 such ratings was less than 0.33, and 95% confidence-inter-

[2] *Webster's New World Dictionary* (College Edition), 1953.

val around any such average was less than ± 0.7. Now, on this basis alone, the two largest differences given in Table I attract our attention since the confidence intervals for these means do not even come close to overlapping.

We are dealing here, however, with the average difference between 36 paired ratings on a given scale. The standard error of this average difference will be less than $0.47\ (1 - r)^{\frac{1}{2}}$, where r is the correlation among pairs over the 36 *S*s. Even if $r = 0$, we would consider the three largest differences "statistically significant." If r is as much as 0.5, a not unreasonable assumption, then the standard error of an average difference is going to be less than 0.33 and we are going to be seriously interested in at least the eight largest differences.

The moral of these observations on the Wiener and Ehrlich paper is: first, that differences are as important as similarities, particularly when highly similar objects are under study; and secondly, that the value of statistics reported without estimates of their precision is seriously impaired.

The semantic differential has much to offer the research social scientist in bringing out differences in meaning between apparently similar concepts. This is the line that should be stressed when measurements of meaning are reported.

University of Chicago BENJAMIN WRIGHT

HUSBAND AND FATHER-IN-LAW—A REVERSIBLE FIGURE

In 1930 Boring introduced W. E. Hill's reversible figure "my wife and my mother-in-law"[1] to psychologists. This figure has since appeared in a variety of textbooks, and it has been used in numerous experimental studies.[2] Its use as a reversible figure for studies in perception is limited by the fact that the wife is seen more frequently than the mother-in-law. In Leeper's study, the wife was reported 65% of the time.[3] In our own study, it was reported 78% of the time by a sample of young adult men and by 94% of a sample of elderly men.[4]

A new figure with masculine motif, "husband and father-in-law," which

---

[1] E. G. Boring, A new ambiguous figure, this JOURNAL, 42, 1930, 444-445.

[2] R. W. Leeper, A study of a neglected portion of the field of learning: The development of sensory organization, *J. genet. Psychol.*, 46, 1935, 41-75; V. R. Carlson, Satiation in a reversible perspective figure, *J. exp. Psychol.*, 45, 1953, 442-448; Jack Botwinick, J. S. Robbin, and J. F. Brinley, Reorganization of perceptions with age, *J. Geront.*, 14, 1959, 85-88; William Epstein and Irvin Rock, Perceptual set as an artifact of recency, this JOURNAL, 73, 1960, 214-228.

[3] Leeper, *op. cit.*, 41-75.

[4] Botwinick, Robbin, and Brinley, *op. cit.*, 85-88.

complements Hill's "wife and mother-in-law," is shown in Fig. 1.[5] *A* is the ambiguous figure, *B* the husband, and *C* the father-in-law, the alternative perceptions.

FIG. 1. "MY HUSBAND AND MY FATHER-IN-LAW"
A shows the balanced figure; the 'husband' is stressed in *B* and the 'father-in-law' in *C*.

Balance of the alternative versions was tested in the manner described in an earlier paper for Hill's figure.[6] The *S*s were 21 women and 30 men, volunteers ranging in age from 18 to 26 yr., and in formal education from 12 to 16 yr. They were drawn from a pool of workers quartered in the Clinical Center of the National Institute of Mental Health. All apparently were in good health.

When presented with the reversible figure (Fig. 1A), 10 of 21 women, and 14 of 30 men reported the 'husband,' while the other *S*s reported the 'father-in-law.' These data clearly suggest the nearly equal likelihood of the two percepts. With verbal suggestion that "another picture could be seen," only two women and five men reported the alternate percept, but, after presentation of the appropriate, less ambiguous figure (Figs. B or C) only three women and four men *failed* to report the alternate percept when once again shown the reversible figure. No significant difference was found between men and women *S*s in the perception of the figure.

JACK BOTWINICK
National Institute of Mental Health

FIFTY-SEVENTH ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The fifty-seventh annual meeting of the Society of Experimental Psychologists was held at Princeton University on March 31 and April 1, 1961.

---

[5] We thank Mr. George P. Marsden, Chief, Medical Arts Section, DRS, N.I.H., for drawing the figures.

[6] Botwinick, Robbin, and Brinley, *op. cit.*, 85-88. We thank Mr. Joseph S. Robbin for his assistance in the collection and tabulation of data.

R. M. Gagne, Chairman of the Society for the year, presided at the business meeting and at sessions for the presentation of scientific papers.

Members and fellows present were: Bartley, Bray, Carmichael, Fitts, Gagne, Garner, Geldard, E. J. Gibson, J. J. Gibson, Graham, Grant, Harlow, Humphreys, Hunt, Hurvich, Irwin, Kennedy, Licklider, Melton, Miles, N. E. Miller, Morgan, Mueller, Neff, Newman, Postman, Pratt, Riggs, Rosenblith, Schlosberg, Solomon, Spence, Stellar, Teuber, Wendt, and Wickens.

James E. Deese, Robert Galambos, Austin H. Riesen, and Donald R. Meyer were elected to membership in the Society, bringing the total membership to 64 members and 20 fellows.

Members presented reports of research in progress at scientific sessions held in Eno Hall the afternoon of the 31st and the morning of the 1st.

The Warren Medal for 1961 was awarded to Carl I. Hovland "for his systematic analyses of human communication as applied to four areas of research—verbal learning, conditioning, concept formation and attitude change."

The Society accepted the invitation of the University of Michigan to meet in Ann Arbor in 1962. Arthur W. Melton was elected Chairman for 1961-62.

Princeton University JOHN L. KENNEDY

FIFTY-THIRD ANNUAL MEETING OF THE SOUTHERN SOCIETY FOR PHILOSOPHY AND PSYCHOLOGY

The Southern Society for Philosophy and Psychology held its fifty-third annual meeting at the Atlanta-Biltmore Hotel in Atlanta, Georgia, on March 30, 31 and April 1, 1961. The Departments of Psychology and Philosophy of Emory University were hosts.

Twenty-nine papers were presented in Philosophy, and thirty-four in Psychology. The special award papers were presented by William T. Blackstone (University of Florida), on "Are metaethical theories normative?", and by Dempsey F. Pennington (University of Alabama), on "Parameters of interference-times in rats as a function of degree of brain damage." These are selected annually from among those papers presented by members who have not yet received their Ph.D. degree, or who have received them within the past five years.

At the first joint session the "History of the Southern Society for Philosophy and Psychology" was presented by Marjorie S. Harris (Randolph-Macon Woman's College). Wilse B. Webb (University of Florida) served

as Chairman of the second joint session, at which Mayor William B. Hartsfield of Atlanta, Georgia, welcomed the Society to Atlanta, and Rubin Gotesky (Northern Illinois University), President of the Society, gave the presidential address entitled "Science as the scientific institution: A new approach." The third joint session was chaired by Rolland H. Waters (University of Florida). A paper entitled "Some current problems in the study of language," was presented by Charles S. Morris (University of Florida), with responses to Dr. Morris' paper from Charles N. Cofer (New York University), and James K. Feibleman (Tulane University). General discussion from the floor followed.

Eighteen psychologists and nineteen philosophers were elected to membership at the annual business meeting.

The following officers were elected for 1961-1962:

William M. Hinton (Washington and Lee), President; Leroy E. Loemker (Emory University), President-elect; Leland E. Thune (Vanderbilt University), Treasurer. Dan R. Kenshalo (Florida State University, is in his third year of a three-year term as Secretary. Charles Hartshorne (Emory University), and Paul S. Siegel (University of Alabama), were elected as Council Members from Philosophy and Psychology, respectively.

The membership accepted the invitation of the Navail Air Station and Southwestern at Memphis, to hold its next annual meeting in Memphis, Tennessee, on April 19, 20 and 21, 1962.

DAN R. KENSHALO

Florida State University

THE 1961 MEETING OF THE NATIONAL ACADEMY OF SCIENCES

The National Academy of Sciences held its ninety-eighth annual meeting at the home of the Academy in Washington, April 24-26, 1961.

The principal program dealt with frontier research on the utilization of solar energy. There were this year no peculiarly psychological papers.

The Academy announced the deaths of twenty-one members and Foreign Associates, and elected thirty-five new members and four new Foreign Associates. The psychologist, Lorrin A. Riggs of Brown University, was elected to membership. The death of Carl I. Hovland, elected in 1960, was announced. Thus the tale of psychologists remains at twenty-four. Statistically oriented psychologists will be interested also in the election of John W. Tukey of Princeton.

The President of the United States, John F. Kennedy, addressed the business meeting briefly, stressing the dependence of the Government on

science and scientists and its constant need for sophisticated advice from men in position to provide the crucial information on which the Government can base its decisions.

The ten psychologists in attendance were F. A. Beach, E. G. Boring, Leonard Carmichael, Robert Galambos, J. P. Guilford, Heinrich Klüver, D. B. Lindsley, W. R. Miles, N. E. Miller, and Carl Pfaffmann.

Harvard University EDWIN G. BORING

THE 1961 MEETING OF THE AMERICAN PHILOSOPHICAL SOCIETY

The annual general meeting of the American Philosophical Society was held in the Hall of the Society on April 20-22, 1961.

The Society was addressed by Curt P. Richter on the subject of biological clocks. The Penrose Lecture was given by Senator J. William Fulbright on the need for a revival of moral idealism in the United States if free government is to survive. The address at the annual dinner was by Caryl P. Haskins, President of the Carnegie Institution of Washington, who reviewed the crucial role of learned societies in the continuing support of scholarships in spite of political change.

The third Karl Spencer Lashley Award "in recognition of useful and significant work in the field of neurobiology" was made to Lord Adrian of Cambridge, England. The presentation was made by William H. Taliaferro. The first two recipients were Rafael Lorente de Nó and Heinrich Klüver.

Of the 19 persons elected to resident membership, those most closely related to psychology were Frank A. Beach of the University of California at Berkeley and Frederick Mosteller of Harvard University.

The psychologists present at the meeting were Leonard Carmichael, Walter R. Miles, Curt P. Richter, and Edwin G. Boring. Among the six members who had died during the year, the youngest was Carl I. Hovland.

Harvard University EDWIN G. BORING

AN ACKNOWLEDGMENT

The JOURNAL is indebted to Professor John F. Shepard, a former student and long-time colleague of Professor Pillsbury, for the portrait and signature reproduced in the frontispiece of this number. Both are dated in the mid-1940s when Pillsbury was about seventy-five years old. The portrait is the last one Professor Pillsbury had taken.

K.M.D.

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

The Human Side of Enterprise. By DOUGLAS MCGREGOR. New York, McGraw-Hill Book Co., 1960. Pp. x, 246. $4.95.

There have been so many dull books telling what psychologists do in the service of management that it is a most refreshing contrast to find a psychologist who applies his technical insights to an incisive analysis of managerial behavior. Douglas McGregor, whose career includes both ample psychological research and high-level administrative responsibility, here offers to management a devastating critique of many traditional practices and a stimulating way of thinking about the proper role of the manager.

Outstanding from my point of view are the excellent uses made of the concepts of perception and motivation. This does not mean that either term is often mentioned. Rather the mode of conceptual analysis is such that one rarely loses sight of the importance of these two decisive features of worker and of managerial behavior. We must constantly consider how a policy will be perceived, and what motives it will satisfy.

Every managerial act rests on a theory about human nature. Interpretations of worker coöperation and noncoöperation involve unconscious, unstated assumptions about motives. Few managers are receptive, however, to the most sophisticated psychological theorizing about motivation; each is sure that his own theory is more accurate. His rejection is often based on ego-motivation; acceptance of a new theory "threatens some cherished illusions" (p. 8). He has, moreover, a built-in defense against failure; "when people respond to managerial decisions in undesired ways, the normal response is to blame them" (p. 10). Thus the manager, despite repeated failures to predict worker behavior, clings to his theory.

Conventional management relies upon what McGregor calls "Theory X," which involves assumptions of human laziness, the necessity of coercion, normal preference for dependency, and implicitly, therefore, the status of the manager as a "higher type" who is not lazy, is willing to take responsibility, and has the right to coerce others into doing his bidding. The contrasting "Theory Y" alleges that humans like work as well as play, will exercise self-direction toward objectives they accept, seek responsibility, and have wide potentialities for intelligence and ingenuity.

The inadequacy of Theory X is demonstrated by reference to a number of common problems. Individual incentive plans do not induce added productivity because they run counter to major noneconomic motives, such as the desire for group approval. Performance appraisals, as customarily administered, do not inform the subordinate of his strong and weak points, do not motivate him, but tend to keep him in a state of childlike dependency which is irritating and conducive to obstructive behavior. Setting up a 'staff' for the purpose of checking up on 'line' performance engenders friction and evasive tactics rather than coöperation.

As a rule, management assumes that workers want only economic goals. But wages and fringe benefits lead to enjoyment *only when the worker leaves his work-place.* "It is not surprising, therefore, that for many wage earners *work is perceived as a form of punishment* which is the price to be paid for various kinds of satisfaction away from the job" (p. 40). Managerial tactics ignore the potency of Maslow's "higher needs." Managers, while believing in Adam Smith's view that entrepreneurs seeking individual advantage benefit society, cannot accept the notion that workers so motivated will benefit the company. Hence the emphasis must be on authority, control, and management prerogatives.

Most of these criticisms are not new. Some of them have been more graphically illustrated by Melville Dalton (*Men Who Manage,* Wiley, 1959.) What is particularly impressive about them is the quiet way in which a statement perfectly acceptable to the typical manager is shown to rest upon the assumptions of Theory X, and how inevitably these assumptions lead to friction, noncoöperation, and the development of groups which defeat managerial purposes. After the analysis has been repeated in five or six major areas of management strategy, the impact seems tremendous. Unfortunately, even if we can get managers to read this book, their defense mechanisms will probably be so effective that the meaning will not get through.

Some of McGregor's points also penetrate the hide of the psychologist. He raises some very tough questions about the ethics of using projective devices to get information in the service of managerial goals, not for the benefit of the client. "Management's legitimate concern is with performance. . . . The question is whether management has a right to go behind the performance to the diagnosis of its causes *when those causes are personal and private*" (p. 107). Do we, as consultants, have a right to reveal confidential data to a man's boss regarding these inner weaknesses?

Likewise he derides the notion of using the appraisal interview as a form of counseling or therapy. "To attempt to counsel in a formal appraisal interview," he writes, "is as much a travesty as to attempt bribery of a victim during a holdup" (p. 86). If the superior is in a position to punish his subordinate for failure to conform, any attempt to whitewash such advice as 'counseling' is a fraud.

The book is as good on the positive side as it is sharp on the negative. Examples are offered of good salary administration (college heads would be well-advised to read pp. 90-99 and 194-199), of good staff-line relationships, of good "managerial development." Of the latter he says characteristically, "In the last analysis the individual must develop himself" (p. 191). These examples build on the assumptions of Theory Y: that men have needs beyond the biological-economic level, and indeed that management has been successful enough that these low-level needs, being satiated, are no longer usable as motivators. Men can, however, find objectives for themselves which are integrated with objectives of the company and, under these conditions, maximal productivity and job satisfaction result.

McGregor shares the reviewers feeling that the Scanlon Plan contract is the vehicle through which psychological insight can best be applied to union-management relations. He is, however, inclined to avoid pat solutions to all problems; what he offers is a way of thinking about human nature and the integration of individuals into organizations. Executives who are intrigued by his speeches, he comments, are always trying to find a formula by which participation and integration can be im-

posed on their organizations. It is, however, precisely this pattern of domination from above which is incompatible with Theory Y. The executive can provide a favorable climate for the growth of integration but he cannot compel it. Subordinates can always misperceive offers of coöperation as attempts at subtle manipulation—and sometimes this is veridical perception.

To what extent is the behavior of management (based on Theory X) responsible for our present problems of featherbedding, lack of interest, diversion of effort outside the organization? McGregor apparently agrees with Clinton Golden's comment that "over the long run, management gets the kind of labor relations it deserves" and extends this to staff-line relations, middle-management behavior, and so on. He points out, for example, that "middle and lower management groups tend to develop protective mechanisms which, although more elaborate and considerably more costly to the organization, are psychologically identical to those developed by workers to defeat the administration of individual incentive plans" (p. 150).

The insistence on dependency, on direction and control from the top, keeps subordinates from satisfying needs for self-esteem and self-actualization. It ignores the realistic fact that managers are dependent upon workers for the achievement of goals. Managerial expectations of childish behavior by subordinates may set off the very actions which are feared. Men who conform willingly to the demands of superiors may make very poor leaders for the organization in the future, because the conventional policies drive out the individualist, the person who thinks for himself, so that he is not available to step into a leadership role.

It is easy to see that I am enthusiastic about this book. I could find a few points to criticize, or at least could allege that they should have been handled differently, but any such comments would be out of place because the overall gestalt of the book is so good. I hope that teachers of industrial psychology and personnel administration will immediately start requiring it of their students; and, even more important, that they learn to think about industrial problems in this framework instead of according to the obsolete assumptions of Theory X—which, like Brand X, deserves only the ash-heap.

ROSS STAGNER
Wayne State University

Applications of Information Theory to Psychology: A Summary of Basic Concepts, Methods, and Results. By FRED ATTNEAVE. New York, Henry Holt and Company, 1959. Pp. vii, 120. \$3.75.

In this clear, compact little book, Attneave's main aim is "to impart a basic knowledge of the subject to the reader whose prior acquaintance with it is minimal" (p. vi). To this end, the first three chapters present tools for measuring information transmitted, and, with these, a selective survey of the findings; the last chapter is an extremely succinct treatment of the broader possibilities of application to psychological problems. The text accomplishes its main aims, and it is high time it was written. The level of preparation required of the reader increases toward the end of the book which closes with an 80-item bibliography; an appendix on "The Calculation of Information Measures from Variance Statistics"; a second appendix containing tables of $\log_2 n$ for $n \leqq 1000$ and a $n \log_2 n$ for $n \leqq 500$ and a graph of $p \log_2 1/p$; and an index.

Chapter headings are "Information: from intuitive to quantitative concepts" (12 pp.); "The uncertainty and redundancy of happenings in a sequence" (28 pp.), "Man's ability to transmit information" (38 pp.), "A new approach to some old problems" (8 pp.).

Attneave's first intention "to set forth a comprehensive survey of informational methods useful in psychological research . . . in enough detail to make them available . . . as practical research tools" (p. vi), has been fulfilled successfully, and Attneave is consistently alert to the general problems of psychology.

The intention to discuss "those areas in which the informational concepts have contributed to a reformulation of old problems in terms which suggest new and profitable lines of investigation" (p. vi) has not been fulfilled as satisfactorily. Although Attneave's selection is judicious, his organization creative, and his suggestions for new lines of investigation stimulating, the brevity of the last chapter (and perhaps of the few pages which precede it) interferes with the appreciation of the psychological problems to which these new methods can be applied. Thus, a discussion of Oldfield's ideas on encoding (in 35 lines) can give the reader a bare idea of Oldfield's thinking, but the three lines on the perceptual and cognitive *schemata* of Woodworth and Bartlett (really quite important, in Attneave's view) can provide only an identifying reference. In a new field of research, what has been done may, however, be less important than conveying some idea of what *can* be done—of the areas in which application has not yet been made. In his last chapter, Attneave attempts to sketch the applicable concepts by example; he has, however, simply not used enough space (about a dozen pages, counting the eight pages in the last chapter) for this.

Consider the use of measures of information-theory, which provide an alternative scaling method to the more conventional psychophysics of discrimination (yielding quite different results). This in itself, of course, constitutes psychological (and engineering) application but, without a general programmatic psychological purpose, we don't know how to extend it to other problems. Aside from scaling for scaling's sake, one major aim of the traditional search for the *jnd* was to determine the "elements of the mind"; *bits,* which are dimensionless and lack *attributive quality,* cannot substitute for the *jnd* in this sense (*i.e.,* we can analyze 'experience' into *bits* as well as into *jnds*, but we could not reconstitute it from the former as was once thought possible with the latter). As a second use, *bits* may provide units for some form of decision-theory model, but then decision-theory would provide the framework for application, not information-theory. Finally, we may build 'subjective scales' to use as tools in other areas: scales of *identifiability* built from *jnd*s seem to be invalid when tested against informational measurements, and the *bit* clearly has greater 'face-validity' and hetermodal generality than the *jnd* as a unit of identifiability. 'Identifiability' as a construct extends some unspecified distance beyond information-measurement, and very little in the way of actual validation is discussed in this book. All *jnd*s were also once assumed equal, and the cross-dimensional comparability of the *bit* for psychological purposes is really an empirical, not a mathematical question.)

There is clearly more 'psychology' in 'information-theory' than these measures which are "as completely neutral with respect to psychological schools and controversies as is analysis of variance or χ^2" (p. 81), otherwise the widespread inter-

est is inexplicable: "it would be quite unrealistic to deny that certain broad concepts of information theory, which have little or nothing to do with specific statistical techniques, are also having their effect on psychological thinking. . . . The concepts of information theory are peculiarly compatible with certain points of view in psychology" (p. 81). If this is true, and I believe it is, we should be given a more explicit discussion of such transferrable broad concepts and of what occasions such compatibility.

Is it simply the central place held by probability mathematics, as in the approaches of Brunswik, Binder, Estes, Bush, etc., that makes for compatibility? I do not think so, since Attneave includes the hypothetical nerve-nets of Pitts and McCulloch, which are not at all probabilistic in the same sense. Is it the use of computer analogies to human behavior? "That the nervous system is a complex communications network, that the brain functions as a kind of computer. . . . Although many psychologists . . . argue that other approaches . . . are more productive, few would [hold] . . . that this point of view is untenable or erroneous" (p. 42). It seems to me, however, that this is really only an up-to-date version of one article of Western feeling, that "man is 'only' a machine"; it may have emotional appeal, but no usable content until the specific *kind* of machine is spelled out—and that leaves us with the very question whose answer we seek.

What to me seems to comprise the heart of the psychological application of information theory is that, without invoking any physical or physiological models at all, it "offers a methodology for quantifying *organizations* and *patterning*, . . . organization is demonstrably measurable in informational terms: roughly speaking, organization and redundancy are the same" (p. 82), and these measures can be applied both to stimulation and response, to 'input' and 'output.' *Much of the theoretical superstructure which has encumbered psychology, has been erected because of the absence of such techniques.* Earlier associationist attempts to deal with patterns of stimulus-variables, or resultant responses, broke these into associations between local patches of stimulation and local fragments of experience because there were no tools to measure organization, nor any plausible physiological mechanisms by which to explain it. Similarly, *Gestalt* theory attempted to talk about "laws of grouping" (only now being pushed into some form of testability) and speculative physiological analyses of organization (via "psychophysiological isomorphism," out of a mixture of physical chemistry and primitive network-analogies). With today's informational approach, "Gestalt psychologists must now be prepared to see their most cherished principles subjected to experimental tests . . . and perhaps modified as a result" (p. 82), and *organization* can be dealt with without either denying its existence, or retreating into speculative explanatory internal mechanisms. Toward this end, information-theory has surely made a beginning; its present tools have not been fully exploited, and other tools lightly sketched by Attneave (such as the measures of "topological information content" of Rashevsky and of Trucco, McKay's "Logon content," and the Pollack-Klemmer "coordinality") promise more powerful application to more 'natural' stimulus-variables.

Although I wish that the final chapter had been twice or three times its published length, this is an excellent book, which fulfills most of its intentions and meets an acute need.

JULIAN HOCHBERG
Cornell University

An Anatomy of Leadership: Princes, Heroes, and Supermen. By EUGENE E. JENNINGS. New York, Harper & Bros., 1960. Pp. xvi, 256. $3.75.

This book is written in the same easily comprehended, popular style of the *Lonely Crowd* and the *Organization Man.* Like these, the book is a commentary upon the current American scene. Yet, it is more. For it contains the best brief summary available of psychoanalytic and philosophical thinking on leadership. But inevitably, this discourse forces the author to devote much space to the situation-person controversy which this reviewer regards as a metaphysical issue since either or both can be shown to vary in importance from one circumstance to another.

The author sacrifices conceptual utility and elegance for conventional, and therefore more easily understood, labels. Using brief sketches of tycoons of the past, the book becomes a polemic against bigness, conformity, teamwork, and consumption. There are good guys and bad guys. Leaders are good guys; executives are bad guys. A leader is a heroic, persuasive, risk-taking innovator who is willing to attempt and ultimately succeeds in getting other people to follow him for their own good. This hero may turn into a superman who goes beyond rational and empirical data depending on faith, action, and intuition. Current heroes in American society are likely to be heroes of leisure and consumption.

The bad guy, the executive, is the efficient executor of plans already developed. He is the self-oriented prince whose goals center in the struggle for increasing his own power. While he is now less likely to use coercion to maintain himself in office, he still maintains his power in the guise of the team captain who ostensibly permits the team to make the decisions. He is more subtle in manipulating people and involves himself in less risk.

Jennings' list of characteristics describing his superior man (the very good guy) following a flippant review about the decline and fall of the summarizers of the traits of leadership, brought to mind images of noble college faculty orators. The image was fixated later by: "perhaps once every 5 to 7 years (the superior man) should be given a year off with pay so that he can read and study and perhaps even write" (p. 239).

Sociological writing today is reminiscent of the classic literature of Ancient Greece which looked wistfully back to a golden age of heroes, of task-oriented demigods, of supermen who knew what was Right, who had an absolute sense of conviction in what was the good, when men were men, boys were boys and goals were clear. Today we live in an age of skepticism, specialization, adjustment, consumption, and conformity. New leadership in production is less required in the western world than new leadership in solving our social problems and our problems of consumption. The abilities of these new leaders even in production may be very different from the cunning princes or visionary heroes of the past.

Looking at past behavior of leaders may provide few clues as to what is to be demanded in this new world of technological advance in communication and organization since 1950. We may need to design radically new ways of organizing groups of men (and computers?) of diverse technical competence speaking different technical languages to work coöperatively and relate to each other rather than bolster our potential hero-princes. The development of the atomic submarine or of Polaris involved not only heroes like Rickover but new concepts of organization.

Social scientists concerned with the current American scene, are victimized by

their own projections and value judgments. The sociologist seems unable to free himself from the role of participant despite his desire to be a dispassionate observer. Yet, these commentators write with more conviction and certitude than the typical experimentalist. Unfortunately for the commentators, organizations differ from loose confederacies to highly centralized one-man companies. Executives differ also from being Jennings-like heroes of high culture, sensitivity, brilliance, and perspective to being the highly selforiented and narrow minded bad guys. The Whytes, Riesmans and Jennings generalize from small samples. Often, their main propositions are drawn from the writings of contemporary popular literature written by a highly select group of observer-interpreters with their own axes to grind.

Although most needed now is the application of controlled experimentation to the study of leadership and organization, and despite my misgivings about the conclusions of the author, I found it refreshing to read such a scholarly presentation. Moreover, commentators such as Jennings are needed in scientific work to inject issues of values and ideals where the more pedestrian amongst us fear to make our own wishes; norms and personal beliefs known. In a sense with this book, Jennings adds his own name to the long list of political philosophers beginning with Plato whose views he well summarized and expanded upon.

Louisiana State University BERNARD M. BASS

Motivation: A Systematic Reinterpretation. By DALBIR BINDRA. New York, The Ronald Press, 1959. Pp. iii, 361. $5.50.

We owe thanks to Dalbir Bindra for this interesting discussion of some causes of some motivational phenomena. He could easily be criticized for what he did not attempt to do or for what he did not achieve, but such criticism might hide the clarity that students of all levels can find in this book. Let us look at the achievement.

The achievement is a victory of theoretical and empirical restraint in a time of the broad romantic sweep in the study of things motivational that is represented in the Nebraska Symposia, the regrowth of genetic thinking in behavior analysis, expansion of personality and social psychologies and the confusing breakdown of methodological canons as summarized by Koch in the third volume of the Project A report. Those who feel overfed by such richnesses may find in Bindra's book a temporary antidote.

The distractions listed above have not kept Bindra from his task. His task is to reshuffle the animal experiments on motivation into a conceptual scheme which reflects some of the insights derived from the analysis of instrumental activities. This, in fact, is what "motivational phenomena" *are* to Bindra—important interdependent classes of instrumental acts the interdependence of which have been documented experimentally.

Bindra decides not to argue about responses as movements, but accepts the notion of acts as observable classes of movements which are classed together because they affect the observable world in a similar way. In fact, they usually are involved in seeking a "goal," which is "an incentive that is chosen by the investigator as a reference point for describing observed behavior." He returns to Holman's notion of goal seeking as "imminent in action," and then analyzes this notion into three

dimensions which can be measured. To be goal directed, an act must have some degree of at least one of the following, somewhat orthogonal, properties: "appropriateness," "persistence," and "searching."

There is careful thinking here, and this is an achievement, albeit a limited one. The limit arises, in part, because Bindra's notion of "appropriateness" is only "meaningful" when applied to a sample of behavior which represents repeated trials under conditions of controlled variation in the stimuli connected with the goal. This limit, however, does no harm because Bindra deals only with data that stay within such limits (with the exception of a few Hebbian experimental anecdotes which occasionally serve the ancient role of "experimentum crucis.")

Then, if we are of a mind to ask, "how does behavior become goal-directed?", we will find the ready, though charmingly tentative, answer: repeated or continued exposure to a situation is probably necessary, and it is mighty significant to note that there is a selective strengthening (or reinforcement) of certain of the responses that occur in a matrix of general activity.

Bindra decided to deal with goal-directed activities organized around the novelty aspect of the environment (exploration), all activities organized around a time-interval (general activity), around problems (problem-seeking, including play), around terminating an existing stimulus-situation (withdrawal and aggressive activities), around food and water transfer (eating and drinking), around the "genital function" (sexual activity) and around offspring (maternal activities). Such classes of activities become recognizably "goal-directed" in all animals, they also "seem to emerge from a state of high excitement, lacking direction." Most important, they all display dramatically the control on behavior from internal and external conditions. The book is then dedicated to showing a student (or an expert) how repeated experience and reinforcers help the class to emerge recognizably, and how various kinds of external and internal conditions have similar effects on all these important response classes.

Bindra believes simply in talking about *things* motivational without a word about "motives" or "drives." Part of the pleasure of reading the book is to see how he avoids doing the latter by organizing the book according to the common classes of controlling conditions rather than the motivational classes. This keeps him from talking too long at any one time about any particular property (habit-strength, goal-directedness, level of arousal) of any particular set of motivational activities, thus decreasing the probability of a slip into inherited or other postulated drives.

This is, then, a behavioristic treatment of some motives defined "on the response side" as the Hullians might say, in order that some limited generalizations regarding controlling factors could be induced in a clear manner. It emphasizes the fact that, by behaviorist rules, it is impossible to distinguish operationally (as yet) between a "drive" and a strong, integrated "habit." Sensory cues, arousal states, and blood factors are all considered, along with reinforcers, as independent variables in the development or control of such strong habits.

Of course, for some readers, these may not be the wear in motivational fashion. Even if this is the case, there are some highlights that a responsible teacher or researcher will not want to miss. The most dispassionate and cogent discussion of the action and nature of reinforcers that this writer has seen for some time, a sensible openness to physiological facts and even tentative theoretical speculation (Hebb in one of his roles); a brave attempt to put "functional autonomy" in its place; a

succinct set of suggestions about how to go about achieving a similar set of functional relationships for human motivation.

Although these latter suggestions are cogent, and possibly even useful, they are advanced without much attempt to do other than dismiss human motivational research to date as too much in the hands of "frequency" counters. Here the richness of the present scene comes back to haunt Bindra's last chapter—there is more afoot in the study of human motives than he makes us aware of.

Bindra's book will probably be more useful as a text—and it has all the signs of a good one—than as a stimulator of research. His empirically-based adumbrations of neurophysiological mechanisms may suggest studies, as will his frequent requests for needed experiments in the context of his review. But there are few signs of the fresh yeast one finds in Berlyne's work on human curiosity, McClelland's attempt at a broad integrative sweep, or Whiting and Child's stretching to subsume fresh realms of data. But when, these developments, as well as most of the data in Bindra's book, arose from the age of the "drive doctrine." Bindra has presented one clear alternative that deserves careful perusal and the test of time.

W. W. LAMBERT
Cornell University

Developmental Psychology. By ELIZABETH HURLOCK. Second edition. New York, McGraw-Hill, 1959. Pp. 630. $6.50.

In this revision of her 1953 textbook, Hurlock seeks to give "as complete a picture as possible of the developmental changes of the total life span of a normal human being." She describes the major areas of development of each of eleven stages from conception to old age, with considerable emphasis on the adjustments necessitated by the maturational process. Some consideration is given to possible causes for deviation from what she describes as normal development and normal adjustment.

Hurlock states a clear bias toward the longitudinal method as the best research technique for studying human development, and she leans heavily on data from such studies when they exist. She also emphasizes the importance of the solution of practical problems as a motivating force behind research in this area. These two criteria seem to have been primary ones in her selection of studies, rather than any clear theoretical orientation. In this revised edition, studies are drawn primarily from the literature since 1940, and there is more material concerning the cultural influences on development than in her earlier book.

The two chapters on middle age are new. Hurlock begins with a rather gloomy picture of the period from 40 to 60 as one of the most dreaded stages in human development, resulting, she feels, from a cultural stereotype which views the period as one of mental and physical deterioration accompanying the cessation of reproductive life.

Hurlock describes this as a period when the individual ought to reach the peak of financial and social success, and reap the benefits of the years of preparation and hard work which preceded it. She also, describes it as a period which requires many adjustments, when changes in physical appearance, energy, and changing roles within the family cause re-adjustment, and changing roles in relation to work and the community necessitate a major revision of the self-concept of the individual.

The first of these chapters deals in detail on the physical changes of the period, the second concentrates on the changes in work and family roles and responsibilities. Much of the material on adjustment to these changes seems to be repetitive in the

two chapters. Since interest in this age-group as a separate stage of development has come about rather recently, subject matter for these chapters tends to be descriptive and has been drawn from a variety of statistical, sociological, and medical sources with a sprinkling of experimental psychological and sociological research.

Any text which covers so wide an age-span as this one must of necessity give a rather small sample of material on each of the various stages of development, and psychologists with a theoretical bias will undoubtedly be critical of the two criteria used for selection of the studies. The developmental material (particularly the physical aspects) is well reported and inclusive, but when the author points to studies which offer reasons for deviation from the normal pattern of development or solutions to adjustment problems the evidence often seems more tenuous than the author indicates. There is a quality of 'oughtness' which is more typical of an earlier era of developmental psychology which will be objectionable to the psychologists oriented toward research, who will perhaps believe that researches in this field still do not clearly indicate what the normal adjustment to the various stages ought to be.

Cornell University CLARA P. MELVILLE

Rigidity of Behavior: A Variational Approach to the Effect of Einstellung. By ABRAHAM S. LUCHINS and EDITH HIRSCH LUCHINS. Oregon, University of Oregon Book, 1959. Pp. xi, 623. $10.00.

From time to time a particular concept has captured phychologists' attention, resulting in a multitude of studies, scattered here and there in the literature, offering similar or contradictory evidence and little closure. Thus the Luchins' book comes as a welcome surprise in its attempt to reëxamine the existing knowledge about the concept of rigidity and to reconcile it with the results of their many original and often ingenious experiments in this area.

The authors argue for the greater use of an empirical and "phenomenon-centered" methodology and indeed their work on rigid behavior is a persuasive demonstration of the fruitfulness of this approach. The major part of their experimental work is organized around variations of what the authors call the "extremum principle," namely the manipulation of those variables (*e.g.* task and environmental, instructions, etc.) which might maximize or minimize the degree of rigid behavior expressed. Their careful pursuit of this method has led to the interesting conclusion that rigidity is more often than not a nonlinear function of the conditions which determine it.

Perhaps the most disappointing part of the book is the writers' treatment of the relationship between concrete thinking and rigid behavior. To date, the nature of this relationship has often been obscured by imprecise definitions and a tendency to use these concepts synonymously. Although the Luchins recognize these points and attempt to provide a more adequate conceptualization of these phenomena, they are not successful. Thus no new systematized formulation is offered and the authors revert back to the traditional criteria of classifying levels of concept-formation when they experimentally investigate the relationship between concreteness and rigidity. The inconsistent results add further to the confusion already existent and the writers' conclusion that "conceptual behavior may depend less on the degree of rigidity and more on the particular testing instrument" (p. 162), only begs the question.

Although some interesting conclusions are drawn from the wealth of material presented, one is left somewhat frustrated by the authors' over-zealous appeal to the importance of field conditions in determining the degree to which any phenomenon may occur at any given time. Thus they are loath to acknowledge that some individuals may be described as more rigid than others and when they evaluate the nature of the relationship between rigid behavior and such variables as prejudice, anxiety, etc. they seem content with the conclusion that if you change the conditions, you change the nature of the relationship. While hasty generalizations based on unreliable measuring instruments or limited populations should be avoided, it is unfortunate that the Luchins use so much restraint when they have available the results of so many carefully planned and executed experiments.

Despite these criticisms the writers offer us a comprehensive presentation of the status of the phenomenon of rigid behavior, and leave us not only with considerable closure about it but with many provocative research ideas for the future.

IRENE R. PIERCE
McLean Hospital, Belmont, Massachusetts

Anatomy and Physiology of Speech. By HAROLD M. KAPLAN. New York, McGraw-Hill Book Co., 1960. Pp. ix, 365. $8.50.

Dr. Kaplan, Chairman of the Department of Physiology at Southern Illinois University has taught anatomy and physiology for twenty-five years. The interest represented by this book was developed largely through instruction of students of speech correction. The writing is clear, systematic and quite comprehensive. The point of view brings together a rare combination of structure with function, of the anatomy of speech seen in terms of its physiology. Recent advances in our understanding of the neural foundations of speech are integrated into a chapter which begins with an elementary coverage of the nerve impulse and the reflex arc. The author presents other basic anatomical concepts, such as discussions of the cell and of tissues. Major coverage is given to autonomic and endocrine regulators of speech, respiration, the larynx, pharynx, palate, and structures similarly participating in the processes of phonation, resonation, and articulation. The anatomy of the ear and the physiology of hearing are included along with some basic coverage of hearing measurement and audiometry. A clearly written glossary increases the book's usefulness to students. This should be regarded primarily as a textbook rather than a reference book, though it is sufficiently unique and systematic to be used as a reference by students of speech pathology.

JOSEPH G. SHEEHAN
University of California, Los Angeles

Elementary Statistics. By SIDNEY F. MACK. New York, Henry Holt and Company, 1960. Pp. ix, 198. $4.50.

This book is perhaps the best of a small but growing set of textbooks on elementary statistics written by mathematicians for the non-technical, *i.e.* the mathless student. Although recognizing the obvious, but often repressed, fact that statistics is applied mathematics and requires some of the basic skills of that discipline, the exposition does not make excessive demands of the student. The approach of the mathematician here is entirely a gain as it has produced an unusual clarity and confidence of exposition.

After a helpful introduction the book begins with a "mathematics refresher" which covers the basic concepts and operations which are used throughout the book.

The section is detailed and stoops to real fundamentals; *e.g.* the arithmetic of fractions, the definitions of accuracy, and linear equations and graphs. It is artfully done. The discussion of fractions teaches the permissible operations of algebra; a section on inequalities and absolute value acquaints the student with the definition of a range of values satisfying an equation. Two chapters are devoted to fundamentals of probability and probability distributions; a single chapter treats the normal distribution; and three chapters discuss sampling distributions, statistical estimation and hypothesis testing. Large sample applications are discussed exclusively. Chi-square and the *t*-distribution are covered in a single chapter. The last chapter presents linear correlation and regression. The emphasis here also is upon basic understanding rather than computation. The least squares line is derived without appeal to the calculus.

The emphasis throughout is upon a thorough teaching of the fundamentals. Some of the usual frills, such as endless verbal discussion of descriptive statistics and the computation of *r*, are omitted. The usual experimental design topics are not covered. Discussion of sampling distributions and hypothesis testing are restricted to the single-sample case. The distribution of differences is not treated. Computational methods are derived in an appendix along with derivations of the mean and standard deviation of the binomial variate and of standard scores.

Some restriction in the range of topics covered and the lack of cute verbal exposition may handicap this book in the soft disciplines. This would be entirely regrettable. Students should derive from the book little of the usual jargon and computational mechanics so quickly forgotten and should achieve some real understanding of the simple and precise notions hidden in the usual study of elementary statistics.

University of Illinois HAROLD W. HAKE

Letters of Sigmund Freud. Selected and edited by ERNST L. FREUD. Translated by Tania and James Stern. New York, Basic Books, Inc., 1960. Pp 470. $7.50.

Sigmund Freud's son has selected a small sample of 315 out of the many thousands of letters that Freud wrote. He confined his selection, as he states, "to letters of a more personal kind in the hope that those who know Sigmund Freud only from his work shall receive here a portrait of the man, of the percipient, thinking, battling human being."

The chief criticism the reviewer has to make about this selection is that it leaves out some of the more dramatic and revealing of Freud's letters. This one can infer from comparing the sample of seven letters to Fleiss with the previously published collection of the Fleiss' correspondence. The letters selected are quite conventional and give no picture of the dramatic relationship that existed between these men. In spite of this qualification, it is worthwhile to have even this small addition to the published letters of Freud. It is hoped that many more of the thousands that are still held from publications will some day be made public.

Washington University Medical School EDWIN F. GILDEA

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

ABERCROMBIE, M. L. JOHNSON, *The Anatomy of Judgment*, New York, Basic Books, Inc., 1960, pp. 156, $4.50.

ABRAMSON, HAROLD A., (ed.), *The Use of LSD in Psychotherapy: Transaction of a Conference on d-Lysergic Acid Diethylamide (LSD-25)*, New York, Josiah Macy Foundation, 1960, pp. 304, $5.00.

ACADEMY OF RELIGION AND MENTAL HEALTH, *Religion, Science and Mental Health, Proceedings of the First Academy Symposium on Interdiscipline Responsibility for Mental Health—A religious and Scientific Concern*, 1957, New York, New York University Press, 1959, pp. xvi, 107, $3.00.

ACADEMY OF RELIGION AND MENTAL HEALTH, *Religion in the Developing Personality, Proceedings of the Second Academy Symposium*, 1958, New York, New York University Press, 1960, pp. xiii, 110, $3.00.

ADLER, JOSHUA, *Philosophy of Judaism*, New York, Philosophical Library, 1960, pp. 160, $3.00.

ADLER, KURT A., and DEUTSCH, DANICA, (eds.), *Essays in Individual Psychology, Contemporary Application of Alfred Adler's Theories*, New York, Grove Press, Inc., 1959, pp. xvii, 480, $2.95.

ANDRY, ROBERT G., *Delinquency and Parental Pathology*, London, Methuen and Co., Ltd., 1960, pp. xv, 173, 21s.

APPLE, DORRIAN, (ed.), *Sociological studies of Health and Sickness*, New York, McGraw-Hill Book Company, 1960, pp. x, 350, $7.50.

BARANSKI, LEO J., *Scientific Basis for World Civilization: Unitary Field Theory*, Boston, Massachusetts, Christopher Publishing House, 1960, pp. 653, $7.50.

BARTLETT, LAILE E., *Bright Galaxy: Ten Years of Unitarian Fellowships*, Boston, Massachusetts, Beacon Press, 1960, pp. xv, 255, $3.50.

BAUMGARTEN, FRANZISKA, *Zur Berufswahl Schweizer Arzte*, Bern, Francke Verlag, 1959, pp. 189, S. Fr. 18.

BLOCH, HERBERT A., (ed.), *Crime in America*, New York, Philosophical Library, Inc., 1961, pp. 355, $6.00.

BOHM, E., *Psychodiagnostisches Vademecum*, Bern and Stuttgart, Verlag Hans Huber, 1960, pp. xvi, 166, DM 22.80.

BOYD, DAVID, *How to Discover Your Better Self*, New York, Vantage Press, 1959, pp. 156, $3.50.

BRAZIER, MARY A. B., (ed.), *The Central Nervous System and Behavior: Transactions of the Third Conference*, Feb. 21, 22, 23, 24, 1960, Sponsored by the Josiah Macy, Jr. Foundation and the National Science Foundation, New York, Josiah Macy Jr. Foundation Publications, 1960, pp. 475, $7.50.

BUCKLEY, MICHAEL J., *Morality and the Homosexual*, Westminster, Maryland, The Newman Press, 1960, pp. xxv, 214, $3.50.

CERDA, ENRIQUE, *Psicologia aplicada: Metodos de exploracion de las aptitudes y de la personalidad*, Barcelona, Editorial Herder, 1960, pp. 421, paperbound $4.50, clothbound, $5.50.

CLOWARD, RICHARD A., *et al.*, *Theoretical Studies in Social Organization of the Prison*, Pamphlet 15, New York, Social Science Research Council, 1960, pp. vi, 146, $1.50.

CLOWARD, RICHARD A,. and OHLIN, LLOYD E., *Delinquency and Opportunity: A Theory of Delinquent Gangs*, Glencoe, Illinois, The Free Press, 1960, pp. xi, 220, $4.00.

COOPER, CLARA CHASSELL, *A Comparative Study of Delinquents and Non-Delinquents*, Portsmouth, Ohio, Psychological Service Center Press, 1960, pp. xii, 240, clothbound $6.00, twiltex cover, $5.25.

COVILLE, WALTER J., COSTELLO, TIMOTHY W., and ROUKE, FABIAN L., *Abnormal Psychology: Mental Illness, Types, Causes, and Treatment*, College Outline Series No. 94, New York, Barnes and Noble, 1960, pp. xv, 298, $1.75.

CROW, LESTER D. and ALICE, (eds.), *Readings in Child and Adolescent Psychology,* New York, Longmans, Green and Company, Inc., 1961, pp. xii, 592, $3.95.

CROW, LESTER D., CROW, ALICE, and MURRAY, WALTER, (eds.), *Teaching in the Elementary School: Readings,* New York, Longmans, Green and Company, Inc., 1961, pp. xii, 571, $3.95.

DUBOIS, CORA, *The People of Alor,* Cambridge, Mass., Harvard University Press, 1960, pp. xxxvi, 654, $10.00.

DUNHAM, H. WARREN, and WEINBERG, S. KIRSON, *The Culture of the State Mental Hospital,* Detroit, Michigan, Wayne State University Press, 1960, pp. xxiii, 284, $5.00.

DUTT, K. GURV, *Existentialism and Indian Thought,* New York, Philosophical Library, 1960, pp. 92, $2.75.

ELMORE, ELLAINE, *Shake Off Your Shackles: Go Free,* Boston, Massachusetts, Christopher Publishing House, 1960, pp. 254, $3.75.

FARNSWORTH, P. R., (ed.), *Annual Review of Psychology, Volume 12,* Palo Alto, California, Annual Reviews, Inc., 1961, pp. ix, 533, $7.00.

FRIED, EDRITA, *The Ego in Love and Sexuality,* New York and London, Grune and Stratton, 1960, pp. viii, 296, $5.50.

FURST, ARNOLD, *A Brief Course in Hyponosis for Salesmen,* San Bernardino, California, AAA Publishing Company, 1960, pp. 120, $3.00.

GLENNY, LYMAN A., *Autonomy of Public Colleges: The Challenge of Coordination,* New York, McGraw-Hill Book Company, 1959, pp. xix, 325, $5.95.

GLOVER, EDWARD, *The Roots of Crime,* New York, International Universities Press, Inc., 1960, pp. xiii, 422, $7.50.

GREENWOOD, DAVID, *The Nature of Science,* New York, Philosophical Library, 1959, pp. xiii, 95, $3.75.

HERBERT, W. L., and JARVIS, F. V., *The Art of Marriage Counseling: A Modern Approach,* New York, Emerson Books, Inc., 1960, pp. 125, $2.75.

HEYWOOD, JEAN S., *Children in Care,* New York, Humanities Press, 1960, pp. x, 229, $5.00.

HILTMANN, HILDEGARD, *Kompendium der Psychodiagnostischen Tests,* Bern and Stuttgart, Verlag Hans Huber, 1960, pp. 252, DM 19.50.

HOEFLIN, RUTH M., *Essentials of Family Living,* New York, John Wiley and Sons, 1960, pp. viii, 282, $5.75.

JOHNSON, DONALD M., *Psychology. A Problem-Solving Approach,* New York, Harper and Brothers, 1961, pp. xiii, 583, $7.00.

JOHNSON, WARREN R., (ed.), *Science and Medicine of Exercise and Sports,* Harper and Brothers, 1960, pp. xv, 740, $12.00.

JONES, MARSHALL R., (ed.), *Nebraska Symposium on Motivation, 1960,* Lincoln, Nebraska, University of Nebraska Press, 1960, pp. xi, 269, paper, $3.25, cloth, $4.25.

JONES, RICHARD M., *An Application of Psychoanalysis to Education,* Springfield, Illinois, Charles C Thomas, 1960, pp. xi, 124, $5.50.

JUNG, C. G. *The Collected Works: III. Volume 3, The Psychogenesis of Mental Disease; VIII. The Structure and Dynamics of the Psyche,* translated by R. F. C. Hull. Bollingen Series XX, New York, Pantheon Books, Inc., 1960, vol. 3, pp. viii, 312, $4.50, vol. 8, pp. x, 596, $6.00.

KANDEL, I. L., *William Chandler Bagley,* New York, Teachers College, Columbia University, 1961, pp. ix, 131, $3.50.

KAPLAN, MAX, *Leisure in America: A Social Inquiry,* New York, John Wiley and sons, Inc., 1960, pp. xii, 350, $7.50.

KING, PETER D., *The Principle of Truth,* New York, Philosophical Library, 1960, pp. 110, $3.75.

LEVITT, MORTON, *Freud and Dewey on the Nature of Man,* New York, Philosophical Library, 1960, pp. 180, $3.75.

LINDGREN, HENRY CLAY, and BYRNE, DONN, *Psychology: An Introduction to the Study of Human Behavior,* New York, John Wiley and Sons, Inc., 1961, pp. xi, 429, $6.50. *Workbook,* to accompany above, 1961, pp. 162. *Instructor's Manual,* to accompany above, 1961, pp. 72.

LOLLI, GIORGIO, *Social Drinking: How To Enjoy Drinking Without Being Hurt By It,* Cleveland, Ohio, World Publishing Company, 1960, pp. 317, $4.50.

MALCOLM, DONALD G., and ROWE, ALAN J., (eds.), *Management Control Systems.* The Proceedings of a Symposium Held at the System Development Corporation, Santa Monica, Calif., July 29-31, 1959, New York, John Wiley and Sons, 1960, pp. xvii, 375, $7.25.

MARTI-ITANEZ, FELIX, (ed.), *Henry E. Sigerist on the History of Medicine,* New York, MD Publications, 1960, pp. xi, 313, $6.75.

MARTIN, JOHN M., *Juvenile Vandalism: A Study of its Nature and Prevention,* Springfield, Illinois, Charles C Thomas, 1961, pp. ix, 145, $6.50.

MCGLOTHLIN, WILLIM J., *Patterns of Professional Education: A Comparative Study of Ten Fields,* New York, G. P. Putnam's Sons, 1960, pp. xviii, 288, $6.75.

MEERLOO, JOOST A. M. *The Dance: From Ritual to Rock and Roll; Ballet to Ballroom,* Philadelphia, Chilton Company, 1960, pp. 152, $4.95.

MESEGUER, P., *The Secret of Dreams,* Westminster, Md., The Newman Press, 1960, pp. viii, 232, $4.75.

MILLEN, JAMES KNOX, *Your Nose Knows,* Los Angeles, Calif., The Cunningham Press, 1960, pp. ix, 166, $3.95.

MOOR, PAUL, *Heilpädagogische Psychologie,* Bern and Stuttgart, Verlag Hans Huber, 1960, pp. 326.

NYQUIST, OLA, *Juvenile Justice,* New York, St. Martins Press, Inc., 1961, pp. xxii, 302, $12.00.

OLYANOVA, NADYA, *The Psychology of Handwriting,* New York, Sterling Publishing Company, Inc., 1960, pp. 224, $3.50.

ORYNSKI, WANDA, (ed.), *Hegel. Highlights: An Annotated Selection,* New York, Philosophical Library, 1960, pp. xvi, 361, $4.75.

PACKARD, VANCE, *The Waste Makers,* New York, David McKay Company, Inc., 1960, pp. x, 340, $4.50.

PEATMAN, JOHN G., and HARTLEY, EUGENE L., (eds.), *Festschrift for Gardner Murphy,* New York, Harper and Bros., 1960, pp. vii, 411, $6.00.

PENNSYLVANIA MENTAL HEALTH, INC., *Mental Health Education: A Critique,* Philadelphia, 1960, pp. xvi, 180.

PHILLIPS, BEEMAN N., DUKE, RALPH L., DEVAULT, M. VERE, *Psychology at Work in the Elementary School Classroom,* New York, Harper and Brothers, 1960, pp. xii, 395, $5.00.

PLUNKETT, RICHARD J., and GORDON, JOHN E., *Epidemiology and Mental Illness,* Monograph No. 6, Joint Commission on Mental Health, New York, Basic Books, Inc., 1960, pp. xvii, 126, $2.75.

REID, D. D., *Epidemiological Methods in the Study of Mental Disorders,* Public Health Papers 2, Geneva, World Health Organization, 1960, pp. 79, $1.00.

ROEMER, MILTON I., (ed.), *Henry E. Sigerist on the Sociology of Medicine,* New York, MD Publications, 1960, pp. x, 397, $6.75.

ROSS, ARTHUR M., and HARTMAN, PAUL T., *Changing Patterns of Industrial Conflict,* New York, Wiley, 1960, pp. x, 220, $6.50.

ROSTOW, W. W., *The Stages of Economic Growth: A Non-Communist Manifesto,* Oxford, Cambridge University Press, 1960, pp. x, 179, $1.45.

ROUCEK, JOSEPH S., (ed.), *Sociology of Crime,* New York, Philosophical Library, Inc., 1961, pp. 551, $10.00.

RUNES, DAGOBERT, *Karl Marx: A World Without Jews,* New York, Philosophical Library, 1959, pp. xii, 51, $2.75.

RUNES, DAGOBERT, *Letters to My Teacher,* New York, Philosophical Library, 1961, pp. 105, $2.75.

SCHAAR, JOHN H., *Escape from Authority: The Perspectives of Erich Fromm,* New York, Basic Books, Inc., 1961, pp. x, 349, $6.50.

SCHAFFNER, BERTRAM, (ed.), *Experimental Aspects of Pediatrics: The Analysis of Behavior in Terms of Control Systems; The Cult as a Condensed Social Process.* Group Processes Transactions of the Fifth Conference, Oct. 12, 13, 14, 15, 1958, New York, Josiah Macy, Jr., Foundation, 1959, pp. 196, $4.50.

SCHOTTSTAEDT, WILLIAM W., *Psychophysiologic Approach in Medical Practice,* Chicago, The Year Book Publishers, 1960, pp. 352, $8.00.

SCHWEITZER, ALBERT, *The Light Within Us,* New York, Philosophical Library, 1959, pp. 58, $2.75.

SEELY, CHARLES S., *Modern Materialism,* New York, Philosophical Library, 1960, pp. 83, $2.50.

SHAGASS, M.D. and PASAMANICK, BENJAMIN, (eds.), *Child Development and Child Psychiatry: A Tribute to Dr. Arnold Gesell in his Eightieth Year,* Psychiatric Research Reports No. 13, New York, American Psychiatric Association, 1960, pp. 225, $2.00.

SIMMEL, GEORG, *Sociology of Religion,* New York, Philosophical Library, 1959, pp. x, 76, $3.75.

SIMPSON, GEORGE, *People in Families: Sociology, Psychoanalysis and the American Family,* New York, Thomas Y. Crowell Company, 1960, pp. xv, 554, $5.75.

STEINER, LEE R., *Understanding Juvenile Delinquency,* Philadelphia, Chilton Co., 1960, pp. vi, 199, $3.95.

STEINER, RUDOLF, *Friedrich Nietzsche: Fighter for Freedom,* Englewood, New Jersey, Rudolf Steiner Publications, Inc., 1960, pp. 222, $4.75.

SUHR, ELMER G., *The Ancient Mind and Its Heritage,* Vol. 2, New York, Exposition Press, 1960, pp. 307, $5.00.

SUTHERLAND, JOHN D., (ed.), *Psychoanalysis and Contemporary Thought,* New York, Grove Press, 1959, pp. 149, $3.50.

TEICHER, MORTON I., *Windigo Psychosis: A Study of a Relationship between Belief and Behavior among the Indians of Northern Canada,* Proceedings of the 1960 Meeting of the American Ethnological Society, Seattle, Washington, University of Washington Press, 1961, pp. xiv, 129, $3.50.

TELFER, DARIEL, *The Caretakers,* New York, New American Library, 1960, pp. 352, 75¢.

TOURNIER, PAUL, *A Doctor's Casebook in the Light of the Bible,* New York, Harper and Brothers, 1960, pp. 256, $3.50.

TRASLER, GORDON, *In Place of Parents: A Study of Foster Care,* New York, The Humanities Press, 1960, pp. viii, 248, $5.00.

VALLON, MICHEL ALEXANDER, *An Apostle of Freedom: Life and Teachings of Nicholas Berdyaev,* New York, Philosophical Library, Inc., 1960, pp. 370, $6.00.

VARVARO, PAOLO, *Pensiero e Vita: Lineamenti di una Fisiologia dello Spirito,* Palermo, G. Mori e Figli, 1960, pp. 387.

WEST, LOUIS JOLYON, and GREENBLATT, MILTON, (eds.), *Explorations in the Physiology of Emotions,* Psychiatric Research Reports No. 12, New York, American Psychiatric Association, 1960, pp. 271, $2.00.

WESTWOOD, GORDON, *A Minority: A Report on the Life of the Male Homosexual in Great Britain,* New York, Longmans, Green and Company, 1960, pp. xii, 216, $7.00.

WHEELIS, ALLEN, *The Seeker: A Psychoanalyst's Search for his Own Life's Meaning,* New York, Random House, 1960, pp. 242, $3.95.

WILLMOTT, PETER, and YOUNG, MICHAEL, *Family and Class in a London Suburb,* New York, The Humanities Press, 1961, pp. xiii, 187, $4.50.

WINN, RALPH B., (ed.), *A Dictionary of Existentialism,* New York, Philosophical Library, 1960, pp. 122, $3.75.

WOOD, ROBERT W., *Christ and the Homosexual,* New York, Vantage Press, 1960, pp. 221, $3.95.

WYMAN, MARY A., *The Lure for Feeling in the Creative Process,* New York, Philosophical Library, 1960, pp. xii, 192, $4.75.

ZULLIGER, HANS, *The Behn-Rorschach Test,* Bern, Verlag Hans Huber, 1956, pp. 200.

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. Stanley Hall

Vol. LXXIV SEPTEMBER, 1961 No. 3

THE KNOWN-SIZE-APPARENT-DISTANCE HYPOTHESIS

By William Epstein, University of Kansas

Most textbook-discussions of space perception include an enumeration of the cues to apparent distance. Among these cues, 'known size' usually is represented as a powerful determinant of static localization under conditions of monocular observation. Discrete changes in the size of the retinal image of an object whose known size remains constant are perceived as corresponding changes in the apparent distance of that object. The identified object is localized at the point in space at which an object of physical size equal to the known size would have to be located in order to yield the given retinal size.

An ingenious demonstration of this effect has been reported by Ittelson.[1] In one experiment, three playing cards were presented singly to *O* under conditions of complete reduction. Each of the cards was placed at the same physical distance from *O*. The task for *O* was to adjust a comparison-stimulus of familiar size which was presented separately until the comparison-object and the standard playing card appeared to be at the same distance. The neat turn in the experiment concerns the sizes of the three cards: one was a normal-sized card, all the dimensions of another one were doubled, while the dimensions of the third card were halved. In this situation, the only cue available for the estimation of distance was retinal size which varies directly with changes in physical size when distance is constant. When known size is invariant, these changes in retinal size ought to be perceived as changes in distance and not as changes in size. The larger card should be localized at a point halfway between *O* and the distance at which the normal card is perceived, and the smaller card should be localized at twice the distance of the normal card. The results for five *O*s confirmed these expectations almost exactly.[2]

This experiment often has been cited in the psychological literature as an impressive demonstration of the effectiveness of the known-size cue. It also has been sub-

* Received for publication May 20, 1960. This study was supported by Grant M-4153 from the Public Health Service. The author is indebted to Mrs. Donna Hardy for her assistance in performing the experiments reported in this paper.

[1] W. H. Ittelson, Size as a cue to distance: Static localization, this Journal, 64, 1951, 54-67.

[2] *Ibid.*, 64.

jected to considerable criticism. Hochberg and Hochberg have argued that Ittelson and others have failed to distinguish between familiar size, on the one hand, and the *relative* size of the stimulus-objects on the other hand, (*i.e.* change or difference in size of similar shapes).[3]

For this reason, Hochberg and Hochberg designed an experiment in which familiar size and relative size were separated.[4] Two figures were presented in a two-dimensional, reversible-screen drawing. One panel contained a drawing of a man, and on the other panel a boy of the same size and approximate contour was represented. The question asked was whether the panel with the boy would appear to be nearer more often than the panel containing the man, a result to be expected if familiar size determined apparent localization. The results showed that familiar size was ineffective in this situation.

In a second experiment, the effectiveness of *relative* size was tested. The same procedure was followed with one difference. Whereas the first experiment held relative size constant while familiar size was varied, the second experiment held familiar size constant while varying relative size. Both panels contained drawings of the same boy, but one was a reduced version of the other.[5] Here, relative size would lead to localizing the panel containing the larger boy nearer than the other panel. The results were in agreement with this expectation. These findings led the authors to suggest that there may be a stimulus-bound correlation between retinal size and perceived distance which makes the postulation of unconscious assumptions (about known size) unnecessary.

Ittelson has criticized both the methodology and the interpretation of Hochberg, and Hochberg and they, in turn, have replied to Ittelson's objections.[6] Instead of considering these arguments, it will be more instructive to describe briefly an experiment by Gogel, Hartman, and Harker.[7] The problem for these investigators was to "investigate whether the retinal subtense of a familiar object can act as a determiner of the apparent *absolute* distance of that object from the observer."[8] This study employed a nonvisual method of measuring perceived distance of the object. *O* was asked to throw a dart to the perceived distance without seeing the results of the throw. Since successive throws might lead to a relative judgment of distance, only the response to the object which is *first* perceived was considered in measuring the absolute distance of the object perceived. The stimulus-object was a normal or double-sized playing card, located at a distance of 10 or 20 ft. in a reduced-cue situation.

The distance-responses for the object initially presented did not confirm the expectations which follow from the hypothesis. Not only did the results fail to agree

---

[3] C. B. Hochberg and J. E. Hochberg, Familiar size and the perception of depth, *J. Psychol.*. 34, 1952, 107-114. J. E. Hochberg and Edward McAlister, Relative size vs. familiar size in the perception of represented depth, this JOURNAL, 68, 1955, 294-97.

[4] Hochberg and Hochberg, *op. cit.*, 111.

[5] *Ibid.*, 111.

[6] Ittelson, A note on 'Familiar size and the perception of depth,' *J. Psychol.*, 35, 1953, 235-240; Hochberg and Hochberg, Familiar size and subception in perceived depth, *ibid.*, 36, 1953, 341-345.

[7] W. C. Gogel, B. O. Hartman, and G. S. Harker, The retinal size of a familiar object as a determiner of apparent distance, *Psychol. Monogr.*, 71, 1957, (No. 442), 1-16.

[8] *Ibid.*, 1.

with any precise predictions of apparent localization, *e.g.* the double-sized card at a physical distance of 20 ft. should be localized at 10 ft., but the less stringent prediction, *e.g.* the double-sized card will appear to be nearer than the normal card, was not confirmed. Under these conditions, perceived distance was totally unrelated to retinal size.

When a similar analysis was performed for all of the four reduced-cue situations collectively (*i.e.* the same *O*s in all four situations), partial support was obtained for the hypothesis of apparent size and distance in its less precise formulation. The implication of this finding is clear. The secondary analysis shows only that the perception *relative*-distance, as some function of *relative* retinal subtense, can occur for successively presented stimuli.

It seems fair to say on the basis of this review that the role of known size in the perception of distance is in doubt. The several studies to be reported here were designed to explore the question further. In experiment I, an attempt was made to modify the normal assumptions concerning the size of familiar objects in order to see whether appropriate changes in perceptual experience would result. The second experiment was intended to determine whether relative size of similar objects is a sufficient cue for apparent localization in an analogous situation. The purpose of the third experiment was to determine whether any systematic localization-responses could be observed in the absence of cues of known-size and relative-size.

Experiment I

If the critical factor in demonstrations like those reported by Ittelson is an assumption concerning the size of the perceived object, then changes in this assumption should modify the effect with regard to the object under consideration. There is evidence that this may occur in other instances when assumptions are believed to determine perceptual experience. The distorted room experiments may be taken as an illustration of this phenomenon.[9] When *O* is allowed to perform some purposeful activity during which he learns about the true shape of the room, then the distortions in apparent size originally obtained disappear. Size-constancy is reinstated because the erroneous distance-estimates upon which the size distortions are partially based are corrected as a result of the modification in *O*'s 'assumptive world.'

In the present study, *O* was placed in a situation which seemed very likely to alter radically his notions about the size of playing cards. *O*

[9] F. P. Kilpatrick (ed.), *Human Behavior from the Transactional Point of View*, 1952, 41-55; Melvin Weiner, Perceptual development in a distorted room: A phenomenological study, *Psychol. Monogr.*, 70, 1956, (No. 423), 1-38.

then was required to make judgments of distance under experimental conditions similar to Ittelson's in all important respects.[10] A control group made similar judgments without any prior modifying experience.

Apparatus. The 'Thereness-Thatness Table' was used in all of the experiments reported in this paper. This apparatus is described in detail in the construction-manual for the Ames demonstrations in perception.[11] It is designed to permit *E* to present the standard- and comparison-stimulus-objects in separate fields. One field, which is viewed monocularly, contains only the standard object. The other field, which is viewed binocularly, contains the variable object with which the standard is compared. In the present instance, the distance of the comparison-object from *O* can be varied continuously, or in discrete steps, by moving it along a track. The field containing the standard is in complete darkness. The field containing the variable is dimly illuminated, allowing *O* to see the comparison-object and a series of vertical guideposts along the left side of the alley wall.

Materials. The following stimulus-materials were used:

(a) Five full decks of playing cards were obtained. One was of normal size ($3\frac{1}{2} \times 2\frac{1}{4}$ in.); two were larger than normal (one twice and one three times); and two were smaller than normal (one one-half and the other quarter size).

(b) Five photographs were taken of the 'three of clubs.' One was of normal size ($3\frac{1}{2} \times 2\frac{1}{4}$ in.), two were enlargements (one twice and the other three times the normal size), and two were reductions (one-half and one-quarter of normal size). Except for differences in size, the five photographs were identical. They were mounted and used as the standard cards in the judgments of distance.

(c) An additional set of 16 photographs of the 'three of clubs' ranging in equal steps from $\frac{1}{8}$ normal size to $3\frac{3}{4}$ normal size was prepared. These photographs were mounted on a black cardboard and were arranged in three rows in order of size.

(d) Another set of stimulus-cards consisted of five blank rectangles of white cardboard which were identical in linear dimensions with the five photographs described above. A series of 16 rectangles which paralleled the playing cards also was prepared.

(e) The comparison-object in all the experiments was a normal-size package of cigarettes.

Procedure. Every *O* served individually. The procedure described here was followed for both the experimental and control groups with one exception: the first stage was omitted for the control group.

In the first stage, the experimental *O* was brought into a normally illuminated room and was seated at a table opposite *E*. On the table were the five decks of cards and a low screen which could be placed between *E* and *O* during the ensuing 'game.' Since *O* often played the game quite seriously, it was necessary to allow him to conceal his cards during the game by interposing the screen between his cards and *E*.

O was then informed that he would be engaged in a variant of the familiar matching game called "War." *O*'s attention was called to the five different-sized cards, and *E* distributed six cards from each deck to *O* and to himself. *E* explained that the special rules of the game made card-size the criterial attribute for success. To win a round, a player had to match both the opponent's number and card-size. By pre-arranging the decks it was contrived that the five critical cards (the three of

[10] Ittelson, *op. cit.*, this JOURNAL, 64, 1951, 54-67.

[11] Ittelson, *The Ames Demonstration in Perception*, 1952, 21-24.

clubs) were always included among the cards distributed. The game usually required about 10 min. for completion, and after the game *O* was asked to return all the cards to their respective decks.

The purpose of this game was to give *O* meaningful experience with cards of different sizes, and, in this way, to accomplish a modification of his usual assumption concerning 'normal' card-size and the constancy of the physical size of cards.

O then was blindfolded and led across the hall into a completely dark room. He was seated in the appropriate position at one end of the 'Thereness-Thatness Table,' and the blindfold was removed. The following instructions were then read to the *O*s of both groups.

In front of you on the right you will be shown a playing card [blank rectangle of white cardboard] for 3-sec. which you will be able to see only with your right eye. I am interested in gettting your immediate impression of how far this is from you. On the left you will be shown a cigarette package which you will be able to see with both eyes. This package can be moved toward or away from you by turning this handle. I wish you to adjust the distance of the cigarette pack until it appears to you to be at the same distance as the playing card [blank rectangle of white cardboard].[12]

Every *O* of both groups made 20 settings—one ascending and one descending—using each of the five playing cards and the five blank, white cards. Half in each group were shown the playing cards first, and half the blank cards first. The order of presentation within each category was randomized according to a table of random numbers. Ascending settings were made first in half of the trials; descending first in the other half. The standard stimulus-object—the playing or blank card—was always 5 ft. from *O*'s eyes.

When these settings were completed, the procedure was repeated, but this time *O* was instructed to select a card from the 16-member series which appeared to be the same size as the card seen in the apparatus.[13] To do this *O* had to turn 90° to the left where the mounted series were suspended from a screen, at a distance of 5 ft. from his eyes. This task was performed under the following instructions:

Now you will be shown each of the playing and blank cards again. Each card will be presented for 3-sec. After each presentation I will remove the card and turn on the light. On your left you will see on the screen, 16 different cards. One of these cards is the same size as the card seen in the dark alley. Please tell me as quickly as you can which one it is. That is, tell me which of the cards on the left appears to be the same size as the one you see in the apparatus. I am interested in your immediate first impression.

The experiment was concluded with a prepared series of questions concerning *O*'s experience while making his judgments of distance and size. *O*'s answers were recorded verbatim.

Observers. Forty students in an introductory course in psychology served as *O*s. None had served in an experiment similar to the present one and none knew anything about the objective of this study or about problems in space perception. There were about an equal number of men and women. Twenty *O*s served in the Experi-

[12] About 20% of the *O*s asked to see the first stimulus twice. Only two *O*s made this request for later stimuli.

[13] The *O*s in Ittelson's studies were questioned about apparent size at the end of the experiment. No other effort was made to learn about this factor.

mental Group, and 20 *O*s in the Control Group. They were assigned to one of the two groups in alternating order of their appearance in the laboratory.

Results: (1) The first aspect of the results to be considered is a comparison of the distance-judgments of the two groups. An inspection of the mean judgments of the distance given in Table I shows very little difference between the groups. An analysis of variance confirms this impression ($F = 1.2$, with 1 and 38 *df.*).

(2) A second point of interest is a comparison of the judgments of the distance for playing and blank cards of the same size. A series of successive contrasts at the 10% level, *e.g.* quarter-sized playing card vs. quarter-sized blank card, were performed using the Sheffe test.[14] None of these differences was significant.

TABLE I

MEAN JUDGMENTS (IN FEET AND INCHES) OF APPARENT DISTANCE IN EXPERIMENT I

Group	Stimulus-cards	Height of stimulus-cards (in.)				
		0.875	1.75	3.5*	7.0	10.5
Exper.	playing	8′3.2″	7′3.4″	5′8.0″	4′4.4″	3′ 6.0″
	blank	8′5.6″	7′3.7″	6′3.0″	5′1.2″	3′11.0″
Control	playing	9′3.7″	8′5.8″	5′8.6″	4′3.8″	3′ 4.9″
	blank	9′3.2″	7′8.7″	6′4.4″	5′2.1″	4′ 1.7″

* The height of a standard playing card is 3.5 in.

Both of these findings are in disagreement with the requirements of the hypothesis that knowledge of size affects apparent distance. In the first case, the two groups do not differ despite the modifying experiences undergone by the Experimental Group. In the second case, no difference in judgment of apparent distance was obtained for objects of known size as compared with objects of unknown size (blank rectangles of cardboard).[15]

(3) The Sheffe-test was used again to compare the playing cards of different sizes with each other and the blank cards of different sizes with each other. For this purpose the results from both groups were combined for each of the stimulus-cards. For example, the combined mean of the distance-judgments of the quarter-sized playing cards given by the Experimental and Control *O*s was compared with the combined mean of the

[14] Henry Sheffe, *The Analysis of Variance*, 1959, 66 ff.

[15] An analysis of the judgments made by the 10 control *O*s who judged the blank cards first also confirms this statement, but an attempt might be made to attribute the effects with the blank cards to the assumption of size-identity of similar cards.

half-sized playing card, and so forth. Only two differences between adjacent means with playing cards (the half-sized vs. normal and normal vs. the double-sized) were significant ($p. < 0.10$). Next, the differences between means one step removed from each other, *e.g.* quarter vs. normal or half vs. double size, were analyzed. All of these differences (six contrasts were tested) were significant ($p < 0.01$).

These results fail to confirm the stringent formulation of the hypothesis about known-size, and they are hardly comparable with the findings reported by Ittelson.[16] Nevertheless, the outcome is not entirely incompatible with the less strict statement of the relationship between known size and apparent distance in that the smaller objects tended to be seen as farther away and the larger objects as nearer.

TABLE II

MEAN JUDGMENTS (IN INCHES) OF APPARENT HEIGHT IN EXPERIMENT I

Group	Stimulus-cards	Height of stimulus-cards (in.)*				
		0.875	1.75	3.5*	7.0	10.5
Exper.	playing	1.9	2.7	3.8	5.6	6.2
	blank	1.9	2.5	3.7	5.7	5.7
Control	playing	1.8	2.6	3.8	5.8	6.5
	blank	1.9	2.7	3.8	5.1	6.6

* The height of a standard playing card is 3.5 in.

(4) The mean judgments of apparent size for the various standards are shown in Table II. It is immediately obvious that the cards of different objective size also are perceived to be of different size. The smaller cards are judged to be smaller and the larger, to be larger. The smaller cards are not, however, judged to be as small as they actually are, nor are the larger judged to be as large as they really are. In other words, the smaller cards are overestimated, and the larger underestimated.

This finding makes the hypothesis under consideration even more untenable as an explanation of the differences in apparent distance which were obtained. The hypothesis asserts that when known size (and hence also perceived size) remains constant, any change in the size of the corresponding retinal image will be perceived as a change in distance. In the present experiment, however, differences in apparent distance were reported even though perceived size varied with the size of the retinal image.

[16] Ittelson, *op. cit.*, this JOURNAL, 64, 1951, 64. The results reported by Ittelson are as close to theoretical expectation as psychological data ever are. For example, the half-sized playing card at a physical distance of 7.5 ft. is localized at an apparent distance of 14.99 ft. An apparent distance of 15.0 ft. would have consistituted perfect agreement with the theoretical requirement.

The objectively smaller cards were seen *both* as smaller and more distant; the objectively larger, were judged to be *both* larger and nearer.

(5) The results of the post-experimental interview support much of what has been reported. Of most importance is the unanimous report that the several cards whose distances were judged did not all appear to be the same size. All of the *O*s indicated that the size-matches made in the final stage of the experiment were also good estimates of the sizes of the cards perceived in the earlier part of the experiment. It should be added, however, that almost all *O*s insisted that they had not been concerned with the variable of size during the distance-settings. The *O*s in the Experimental Group were asked what they thought the purpose of the 'card-game' was. None suggested any relationship between the game and the experimental observations which could be construed as signifying that information gained during the game was consciously employed while the distance-judgments were being made.

The foregoing analysis has shown that the observed changes in apparent distance cannot be explained in terms of the hypothesis of known size and distance. The two experiments which are described below explore an alternative explanation of these effects.

Experiment II

Experiment II was performed to determine whether similar results would obtain when only relative retinal size is operative. For figures of the same shape, difference in size may be a cue for the perception of a difference in distance, the larger one appearing to be nearer. For this purpose, stimulus-objects must be used which are not of any particular known size. Furthermore, a set of objects is required which does not encourage the assumption of size-identity, *i.e.* that all of the objects are the same size, although the specific size of any of the particular objects may be unknown.

Materials. Five disks of different diameter were cut from paper of different colors and mounted on corresponding disks of stiff cardboard. The diameters of the disks 0.5, 1, 2, 4, and 6 in., and their colors were brown, blue, purple, green, and gray, respectively. It seems fairly certain that unmarked disks do not have a fixed subjective or assumed size. Furthermore, the use of different color reduces the likelihood of the identity of size.

A series of 16 white circles ranging in diameter from ¼, ½, 1 and thence by ½-in. steps to 7½ in., was mounted on a black cardboard in a manner identical with the earlier procedure. The comparison-object was the cigarette package used in Experiment I.

Procedure. The procedure was similar to that for the Control Group of Experiment I, with references to playing cards deleted from the instructions and new questions added to the post-experimental interview.

Observers. Twenty students, drawn from the same population as those of Experiment I, served as *O*s. Half were men and half were women.

Results. The results are presented in Table III. A three-part analysis of variance yielded a significant F of 18.44 ($p < 0.01$, with 4 and 76 *df.*) for the means of the five stimulus-disks. Duncan's multiple-range test for multiple comparisons among individual means was then used.[17] The results of this analysis show that the means for the two larger disks are significantly different from the means for the two smaller ones. In addition, the mean judgments of distance for both the smaller and the larger disks differ significantly from the mean for the 2-in. disk. An inspection of Table III reveals that all these differences are in the expected direction. The two smaller disks do not differ significantly from each other, and the same is true for the two larger disks.

TABLE III

MEAN JUDGMENTS (IN FEET AND INCHES) OF APPARENT DISTANCE AND SIZE IN EXPERIMENT II

Mean	Physical diameter of disks (in.)				
	0.5	1	2	4	6
Apparent distance	6′10.3″	6′1.6″	4′7.6″	4′1.1″	3′8.8″
SD	2′ 6.79″	1′9.28″	1′3.75″	1′4.97″	1′7.18″
Apparent size	0.5″	1.1″	1.7″	2.4″	3.0″
SD	0.27″	0.78″	0.77″	0.99″	0.96″

The size-judgments show the same effects as in Experiment I. The objectively smaller disks are perceived to be smaller, but they are nonetheless overestimated. The larger disks are perceived as larger, but their actual size is underestimated.

The post-experimental interviews provided corroborative information. All of the *O*s reported that the disks appeared to them as five different objects of different size. All of the *O*s agreed that the disks presented for size-judgment were the same disks shown for distance-judgment, and that there size-matches also corresponded to the appearance of the disks when their distance was being judged. Finally 16 *O*s reported that the disks appeared at different distances when their size was being judged.

The results of Experiment II are very similar to those reported for Experiment I. The difficulties which were encountered in applying the hypothesis of known size to the results of Experiment I are compounded when a similar effort is made to account for the findings of Experiment

[17] A. L. Edwards, *Experimental Design in Psychological Research*, rev. ed., 1960, 136-140. The table of significant 'studentized' ranges was entered with *df.* = 76, and $\alpha = 0.05$.

II in the same way. It is highly improbable that the effects obtained in Experiment II can be attributed to the operation of any assumptions concerning known size or size-identity. A more likely explanation is to be found in the influence of relative size on apparent distance.

Experiment III

In Experiment III both known and relative size were eliminated as possible determinants of apparent distance.

Procedure. Each *O* was shown only one of the disks employed in Experiment II. Five different groups of *O*s participated in this study; each group judging the apparent distance and apparent size of one of the five standard disks. In addition, another minor change in procedure was introduced. Each *O* made four distance-judgments instead of only two. Of these, two were in ascending order, and two in descending order, in the sequence *ADDA* or *DAAD*.

TABLE IV

Mean Judgments (in Feet and Inches) of Apparent Distance and Size in Experiment III

Mean	Physical diameter of disks (in.)				
	0.5	1	2	4	6
Apparent distance	5′4.5″	6′6.0″	4′6.5″	4′10.0″	4′2.6″
SD	1′7.54″	1′7.60″	1′3.75″	1′ 9.54″	1′8.10″
Apparent size	0.6″	1.5″	2.0″	2.6″	3.3″
SD	1.04″	0.95″	0.69″	1.07″	1.31″

Observers. Eighty students from the population sampled previously served as *O*s. They were assigned to one of the five groups in the order of their appearance for the experimental session. Each group contained 16 *O*s about equally divided between men and women.

Results. The results are presented in Table IV. An analysis of variance yielded a significant F of 5.61 ($p < 0.01$ with 4 and 75 *df.*). An analysis of the differences between individual means using Duncan's technique showed, however, only the most extreme difference, *i.e.,* 6 vs. 1 in., to be significant.[18]

A closer inspection of the means in Table IV provides further evidence of the absence of systematic judgments of size-distance. Two reversals will be noticed. The 1-in. disk is judged to be farther back than the ½-in. disk, and the 4-in. disk to be farther back than the 2-in. disk. This inconsistency is easily recognized if (disregarding statistical significance) the disks are arranged in the order of increasing apparent distance—6 in.,

[18] *Ibid.,* 136-140. The table of significant 'studentized' ranges was entered with *df.* = 75 and $\alpha = 0.05$.

2 in., 4 in., ½ in., 1 in. This should be compared with the arrangement in Experiment II—6 in., 4 in., 2 in., 1 in., ½ in.

The pattern of errors in judging size in this experiment is very similar to the results reported for the two earlier studies. Here again the smaller objects are perceived to be smaller, but are overestimated, while the opposite is true for the larger objects.

DISCUSSION

The experiments described in this paper have not supported the hypothesis of known size and apparent distance. In addition, they have contributed evidence which tends to bolster an alternate explanation based on the influence of the relative size of similar shapes. In this regard the results of the present studies contradict the position adopted by Ittelson and by others in the same theoretical camp.[19] The conclusions and findings of the earlier proponents of the thesis of relative size have been confirmed. Nonetheless, a number of problems remain to confront the hypothesis of relative size.

(1) Can it account for all of the reported findings? Particularly problematic is the finding that "discrete changes in the characteristics of the physiological stimulus-pattern resulting in changes in *assumed size* will be perceived as discrete changes in apparent distance even though *retinal size remains constant*."[20] Thus, a half-sized playing card at an objective distance of 7.5 ft. is localized at a distance of about 15 ft., while a match-box of identical size at the same physical distance is judged to be at a distance of about 9 ft.[21] A related finding has been reported by Hastorf.[22] A rectangular or circular area of light was given a "large assumed meaning of size" or a "small assumed meaning of size." That is, the rectangle was called either an envelope or a calling card, and the circle was called either a billiard ball or a ping-pong ball. The size at which the stimulus-object was set, in order to appear at a specific distance, varied when the assumed size attributed to the object was varied by the suggestion of size, *i.e.* by naming the object.

These changes in apparent distance are in agreement with the hypothesis of known size and apparent distance. Furthermore, they do not seem to be amenable to an explanation in terms of relative size since physical size remained unaltered. In view of this fact, any total dismissal of known size as a cue to apparent distance would not be justified.

(2) With regard to the experiments reported in this paper, the major explanatory problem resides in the fact that a small retinal image appeared as a

[19] See for example, A. H. Hastorf, The influence of suggestion on the relation between stimulus size and perceived distance, *J. Psychol.*, 29, 1950, 195-217.

[20] Ittelson, *op. cit.*, this JOURNAL, 1951, 67. (Italics ours.)

[21] *Ibid.*, 64.

[22] Hastorf, *op. cit.*, 195-217. It is possible to argue that Hastorf has only demonstrated a cognitive effect rather than a perceptual effect, and therefore his study should not be considered in the present context.

smaller than normal object which was at the same time more distant, while the larger retinal images produced the converse effect. An acceptable interpretation of the present experiments must account for both aspects of the resultant experience.

At first it may appear that the explanation is to be found in the hypothesis of the invariance of size and distance which relates perceived size to perceived distance. According to this thesis, a given visual angle subtended by an object uniquely determines the ratio of its perceived size to its perceived distance. Thus, if the distance of an object is overestimated, its size must be overestimated, and if its distance is underestimated, its size must be underestimated.

In the present experiments, the influence of relative size produces an underestimation of the distance of the larger object when compared with the objective distance of the stimulus-object and, therefore, the size of the object is underestimated as compared with its physical size. Similarly, as a result of the cue of relative size, the distance of the smaller object is overestimated, and for this reason the object is perceived to be larger than it actually is. In this account, apparent distance would depend on relative size, and apparent size would depend on apparent distance.

This proposed solution is open to criticism on several counts:

(1) There is a growing body of evidence which casts doubt on the generality of the hypothesis of the invariance of size and distance.[23]

(2) The hypothesis has no explanation for the many 'nonfitting' size-distance pairings which were obtained, *i.e.* overestimations of distance coupled with underestimations of size or underestimations of distance coupled with overestimations of size. (In the first experiment there was 249 fitting and 151 non-fitting size-distance pairings; in the second and third experiments there were 107 fitting and 73 non-fitting pairings.) It is difficult to uphold the invariance relationship when it clearly does not apply uniformly to all *O*s and all judgments.

(3) Finally, the results of Experiment III create serious difficulties for the explanation under consideration. According to the hypothesis of invariance, when the values of any two of the variables, *e.g.* visual angle and apparent size, are given, then the value of the remaining variable is already determined. Since the judgments of size obtained in Experiment III show the same pattern of overestimation and underestimation as do the judgments of the first two experiments, one would expect a complementary pattern of distance-judgments. This was not the case. Errors in size-estimation were not accompanied by appropriate errors in distance-judgment.

The outcome of Experiment III suggests an alternate description of the findings reported in this paper. It is quite possible that judgments of distance and size are unrelated. Apparent distance seems to depend on

---

[23] Such evidence can be found in F. P. Kilpatrick and W. H. Ittelson, The size-distance invariance hypothesis, *Psychol. Rev.*, 60, 1953, 223-31; H. E. Gruber, The relation of perceived size to perceived distance, this JOURNAL, 67, 1954, 411-26; The size-distance paradox: A reply to Gilinsky, this JOURNAL, 69, 1956, 469-75; Noel Jenkin and Ray Hyman, Attitude and distance-estimation as variables in size matching, this JOURNAL, 72, 1959, 68-76.

the relative size of the retinal images of successively presented stimuli, while apparent size may be determined by the absolute size of the retinal image which is being cast by the single object under observation.[24] This description is consistent with all of the data which have been reported. It is not controverted by the presence of non-fitting pairings of size-distance since these two judgments are assumed to be independent of each other. This account also eliminates any questions concerning the primacy of one or the other of the two perceptual experiences. Explanations of space perception often proceed as if one experience is primary, *e.g.* distance, while a second experience is dependent on the first, *e.g.* size. In many cases this results in an explanatory circle in which the dependence of size on distance is asserted at one stage, only to have the dependency relationship reversed at a later stage in the explanation.

Regardless of the popularity enjoyed by this type of explanation, there is no *a priori* reason why one percept must always depend on another. Instead, it is very likely that in many cases apparent size and apparent distance are independently and simultaneously determined by different aspects of the proximal stimulus-situation.[25] Of course, there are instances of close correspondence between the two experiences, but even in these cases the possible relationships are many. Apparent size may decrease with increasing distance,[26] it may increase with increasing distance,[27] or it may remain approximately constant with changes in distance.[28] In the same vein, a long line of stereoscopic experiments has failed to establish any definite relationship between changes in apparent size, as a result of changes in tactile-kinesthetic stimulation, and apparent distance.[29] Sometimes changes in apparent size are not accompanied by any clear change

---

[24] Other investigators have shown that, under 'reduced' conditions, changes in the size of the retinal image of an unfamiliar object will be accompanied by proportional changes in apparent size. See A. H. Hastorf and K. S. Way, Apparent size with and without distance cues, *J. gen. Psychol.*, 47, 1952, 181-188.

[25] A similar conclusion is reached by Gruber, *op. cit.*, this JOURNAL, 67, 1954, 411-427.

[26] A. S. Gilinsky, The effect of attitude upon the perception of size, this JOURNAL, 68, 1955, 173-192.

[27] E. L. Chalmers, Monocular and binocular cues in the perception of depth, this JOURNAL, 65, 1952, 143-153; W. M. Smith, A methodological study of size-distance perception, *J. Psychol.*, 35, 1953, 143-53; Noel Jenkin, Effects of varied distance on short-range size judgments, *J. exp. Psychol.*, 54, 1957, 327-331.

[28] J. J. Gibson, Motion picture testing and research report No. 7, *Army Air Force Aviation and Research Reports*, 1947, 200-211.

[29] For one recent illustration see the experiment reported by Herman. Despite the fact that all *O*s in this study reported a reduction in apparent size with an increase in convergence, "some *O*s apparently responded to decrease in apparent size by reporting an increase in apparent distance, while others responded by reporting decrease in apparent distance" (T. G. Herman, The relationship of convergence and elevation changes to judgments of size, *J. exp. Psychol.*, 1954, 48, 208).

in apparent distance and on occasion may even be accompanied by 'paradoxical' distance effects, *e.g.* an object will appear to be smaller with increased accommodation, but, at the same time, it will appear to recede. In view of these possibilities, it can only be misleading to develop hypotheses whose main justification is their logical derivation from the traditional hypothesis of the invariance of size and distance.[30]

SUMMARY

Three experiments were performed to study the role of known size and relative size in the perception of distance.

In Experiment I, two groups of *O*s judged the apparent distance and size of playing cards and blank cardboard rectangles of different sizes. Prior to making these judgments one group of *O*s (experimental) participated in a game which was designed to modify the assumptions which *O* might ordinarily make about the normal size of playing cards and the constancy of the physical size of playing cards. The *O*s in the second (control) group did not undergo any prior treatment.

The main results were as follows: (1) The experimental and control groups did not differ significantly in their judgments of apparent distance. (2) Judgments of the playing cards (of known size) and blank cards (unfamiliar size) did not differ significantly. (3) The judged distances of the playing cards did not meet the quantitative requirements of the 'strong' known-size hypothesis. (4) The cards of varying physical sizes also were perceived to be of different sizes.

Experiment II demonstrated that similar differences in apparent distance (and size) would be obtained when only relative retinal size was operating. For this purpose five cardboard disks of different sizes and colors were used which did not encourage the assumption of known size or of constancy of physical size.

In Experiment III, the same disks were judged again. This time each *O* judged only one of the disks. In this way, the cue of relative size was eliminated. Under these conditions, no systematic differences in apparent distance were observed, although comparable differences in apparent size were obtained.

The conclusion was reached that in experiments of this nature the determinant of apparent distance is not the known size of the objects, but the relative size of objects presented successively.

[30] A general discussion of this hypothesis may be found in R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1954, 480-486.

COMPARISONS AMONG WORD-ASSOCIATION RESPONSES IN ENGLISH, FRENCH, GERMAN, AND ITALIAN

By MARK R. ROSENZWEIG, University of California, Berkeley

It has now become possible to compare directly word-association responses among several languages, since recent studies in France, Germany, and Italy have employed translations of the list of Kent and Rosanoff which has long been used in America. An initial comparison was made between English and French primary responses (*i.e.* the response given most frequently to each stimulus-word), and it was found that for about half the items the primary responses of the two languages were equivalent in meaning.[1] Next, it was found that the greater the frequency of the primary response in one language, the more likely it is to be equivalent in meaning to the corresponding primary response in the other language.[2] From these results it was hypothesized that associative habits tend to be held in common among different language-communities. In the present report, this hypothesis is tested by comparing the primary responses in English, French, German, and Italian, and a smaller amount of material in Navaho.

METHOD

The English primary responses were obtained from the study of Russell and Jenkins.[3] They administered the stimulus-words of Kent and Rosanoff[4] as a paper-and-pencil test to students at the University of Minnesota in 1952. One thousand and eight complete protocols were obtained. No breakdown of *Ss* by sex was given.

For comparison of primary responses of the two sexes in English, reference will be made to the study of Tresselt, Leeds, and Mayzner.[5] They tested individually 114 men and 115 women. The *Ss* were all between 18 and 25 yr. of age; they were of a variety of occupations and from several regions of the United States.

A French translation of the list of Kent and Rosanoff was administered to stu-

* Received for publication March 19, 1960.

[1] M. R. Rosenzweig, Etudes sur l'association des mots, *Année Psychol.*, 57, 1957, 23-32.

[2] Rosenzweig, Comparisons between French and English word association norms, *Amer. Psychologist*, 14, 1959, 363.

[3] W. A. Russell and J. J. Jenkins, *The Complete Minnesota Norms for Responses to 100 Words from the Kent-Rosanoff Word Association Test*, 1954.

[4] G. H. Kent and A. J. Rosanoff, A study of association in insanity, *Amer. J. Insanity*, 67, 1910, 37-96, 317-390.

[5] M. E. Tresselt, D. S. Leeds, and M. S. Mayzner, The Kent-Rosanoff word association: II. A comparison of sex differences in response frequencies, *J. genet. Psychol.*, 87, 1955, 149-153.

dents at the Sorbonne and at Paris *lycées* in 1955–1956.[6] The same procedures were employed as those used by Russell and Jenkins. Two hundred eighty-eight complete protocols were obtained, 184 from women and 104 from men. The responses of the two sexes were tabulated separately. Although the results for the two sexes were quite similar, in the preparation of the composite table of primary responses the frequencies were so weighted that the men counted as much as the women in determining the final results.

A German translation of the list of Kent and Rosanoff was given to students at the University of Würzburg and at high schools in Würzburg and nearby cities by Russell and Meseck.[7] The same procedures were followed as in the study of Russell and Jenkins. Thirty-one complete protocols were obtained from women and 300 from men; the responses of the two sexes were pooled.

An Italian translation of the list of Kent and Rosanoff was administered to *Ss* in Italy by Levi.[8] The *Ss* were chosen to be fairly representative of the Italian population in age and social class, although the majority were students. About two-thirds of the 229 *Ss* were men; responses of the two sexes were pooled. Ninety-seven of the *Ss* were tested in Turin with a paper-and-pencil test such as those used in the other three languages. The 132 *Ss* tested in Rome heard the stimulus-words from a phonograph record and responded on a sheet with numbered blanks. The responses of these two subgroups were generally similar and were pooled to give a single set of norms. Only primary responses that were given by at least 9% of the *Ss* were reported. For 11 stimulus-words the primary response was given by less than 9% of the *Ss*, and these 11 responses are not available. The responses reported were English translations of the Italian responses.

RESULTS

To compare primary responses in the four languages, a table was prepared which included the stimulus-words, the primary responses, and the percentage of *Ss* giving each response. A judgment was then made as to the equivalences of meaning among the responses for each item, and these judgments were entered as the final column of the table. Because of limitations of space the entire table is not given here, but selected items are given as examples in Table I.[9]

It will be noticed that the Italian responses are given in English, since this is the way in which they were reported. Where no primary response was reported in Italian, as for Item 10, a question mark is shown in the

---

[6] Rosenzweig, *op. cit., Année Psychol.*, 57, 1957, 23-32.

[7] W. A. Russell and O. R. Meseck, Der Einfluss der Assoziation auf das Erinnern von Worten in der Deutschen, Französischen, und Englischen Sprache, *Z. exp. angew. Psychol.*, 6, 1959, 191-211.

[8] Mario Levi, An analysis of the influence of two different cultures on responses to the Rosanoff free association test, Unpublished Master's thesis, University of Chicago, 1949. I am grateful to Mr. Levi for having called his study to my attention and having allowed me to use it.

[9] A 6-page table giving the complete data has been deposited with The American Documentation Institute. Order Document No. 6747, remitting $1.25 for 35 mm. microfilm, or $1.25 for photoprints. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

table. The Italian primary response was not reported for Items 10, 17, 43, 45, 53, 57, 58, 83, 90, 93, and 100. (Except for some of these 11 items, all the items specifically mentioned in the text will be found in Table I.)

TABLE I

STIMULUS-WORDS, PRIMARY RESPONSES, AND THEIR PERCENTAGES FOR WORD-ASSOCIATION TESTS IN ENGLISH, FRENCH, GERMAN, AND ITALIAN

Stimulus-words (English, French, German, Italian)	Primary responses, their percentages, and agreements among four languages				
	English	French	German	Italian	E-F-G-I
1. table, table, Tisch, tavola	84 chair	53 chaise	29 Stuhl	23 chair	E-F-G-I
2. dark, sombre, dunkel, scuro	83 light	45 clair	44 hell(e)	18 light	E-F-G-I
3. Music, musique, Musik, musica	18 song‡	16 note‡	9 Tone‡	10 beautiful	F-G
4. sickness, maladie, Krankheit, malattia	38 health	10 santé	15 Gesundheit	13 serious	E-F-G
5. man, homme, Mann, uomo	77 woman‡	66 femme	52 Frau	39 woman	E-F-G-I
6. deep, profond, tief, profondo	32 shallow	12 creux	49 hoch	16 sea	—
7. soft, mou, wiech, morbido	45 hard	39 dur(e)	39 hart	14 hard, pillow	E-F-G-I
8. eating, manger, Essen, mangiare	39 food	39 boire	23 trinken	30 drinking	F-G
9. mountain, montagne, Berg, montagna	27 hill‡	13 plaine	35 Tal	17 high	—
10. house, maison, Haus, casa	25 home	14 toit	14 Hof	?	—
12 mutton, agneau,* Hammelfleisch, montone	37 lamb	15 brebis	7 Rindfleisch, Essen	18 sheep	—
3. woman, femme, Frau, donna	64 man‡	29† homme	40 Mann	22 man	E-F-G-I
9. beautiful, belle, schön, bello	21 ugly	18 femme	27 hässlich	31 ugly	E-G-I
2. sheep, mouton,* Schaf, pecora	20 wool	25 laine	15 Wolle	18 wool	E-F-G-I
0. square, carré, Quadrat, piazza	37 round	30 rond	14 Viereck	10 large	E-F

* The French stimulus-word and responses for Items 12 and 62 were, as explained in the text, interchanged in the tabulation.

† Incorrectly given as 26% in Rosenzweig's paper in 1957 and also, consequently, in Russell and Meseck's.

‡ Singular or plural.

TABLE II

NUMBERS OF AGREEMENTS IN MEANING BETWEEN PRIMARY RESPONSES AMONG PAIRS OF LANGUAGES

A. All Items (100)

	French	German	Italian
English	48	48	(35)
French		45	(36)
German			(36)

B. Items where Italian Response is Known (89):

	French	German	Italian
English	43	43	35
French		41	36
German			36

The final column of the table indicates which responses have been judged to be equivalent in meaning; *e.g.* the notation E-F-G-I for Item 1 means that all four responses are equivalent, while the dash in this column opposite Item 6 means that no response is equivalent to any other. In the rare cases when two responses are tied for primary rank, as are the Italian responses for Item 7, an agreement was counted if either of the two was equivalent to the primary response of another language. When the Italian response is not available, as for Item 10, it is, of course, considered not to be equivalent to any other response.

Agreement between pairs of languages. Table II gives the number of

primary responses judged equivalent in meaning when any pair of languages is compared. The upper half (A) of the table includes all 100 items. The numbers of agreements between English and German and between French and German are the same as those reported by Russell and Meseck. For the comparison between English and German, however, they reported 47 cases of agreement while we find 48. The numbers of agreements found between Italian and the other languages is clearly biased by the omission of 11 responses from the Italian norms. To make an unbiased comparison, the lower half (B) of Table II was prepared. Here only the 89 items were used for which the Italian responses had been reported. Even in Part B it is the Italian responses which show the smallest numbers of agreements with other languages. It should be remembered that the Italian study differed from the other three in using a more varied group of *S*s and in employing an auditory presentation of the stimulus-words for a large proportion of the *S*s. Both of these factors may have made the Italian responses somewhat different from those of the other groups.

The agreement among the four languages is impressive, almost half the comparisons in any pair of languages yielding agreements. As a yardstick, we may note the amount of agreement between two samples of *S*s drawn from the same population. Where the responses of the two sexes have been tabulated separately for a study, we have two samples tested alike and selected alike in every respect but sex. The primary responses of French men and women agree for 75 out of 100 items. The English data of Tresselt, Leeds, and Mayzner[10] show 82 agreements between the primary responses of men and women. There is little evidence of real sex-differences in either study. Rather, as we shall discuss later, the men and women can be considered simply as two samples drawn from the same language-community. For a further comparison, we can note the agreements between the independent English studies of Tresselt, Leeds, and Mayzner and of Russell and Jenkins.[11] For this purpose, the responses of men and women from the former paper have been pooled. This procedure allows us to compare groups that were tested somewhat differently and that differed in occupation and educational status, both groups however including the two sexes. The primary responses in these two studies agree in 82 out of the 100 cases.[12] Thus, the number of agreements between pairs of Western European languages is only about three-fifths as great as the number of agreements between samples drawn from the same language.

For 21 stimulus-words, the responses of all four languages are equivalent in meaning. If only English, French, and German are considered, there are 36 items

---

[10] Tresselt, Leeds, and Mayzner, *op. cit.*, 150-152.

[11] Russell and Jenkins, *op. cit.*

[12] It is worth noting that the primary responses of Tresselt, Leeds, and Mayzner, based on 229 *S*s, agree so closely with those of Russell and Jenkins, based on 1008 *S*s. This fact indicates that a group of about 200 *S*s is adequate to obtain stable primary responses, and it helps to justify our use of the French, German, and Italian studies, each of which employed groups of about this size.

for which all primary responses agree. (Russell and Meseck judge that there are 34 cases in which the three primary responses agree.) Moreover, when the primary responses do not agree, the primary response of one language is often equivalent to the secondary response of the other language. Russell and Meseck have found this relationship in comparing English and German,[13] and the same result had also been reported in the initial comparison of English and French responses.[14]

The degree of similarity among the languages is probably even greater than appears from these results, since any inadequacies of translation of the stimulus-words inevitably lead to differences of response. Thus, Items 12 and 62, 'mutton' and 'sheep', should both be translated as 'mouton' in French. To avoid repeating a stimulus-word, 'agneau' ('lamb' in the sense of 'the young of the sheep') was used for Item 62. The responses to Item 12 showed that 'mouton' was understood on the sense of 'sheep'; therefore 'mouton' and its primary response 'laine' (wool) are tabulated as Item 62. 'Agneau' and its response are placed as Item 12 ('mutton'), and agreement cannot be expected here. In the case of Item 80, 'square,' the English responses show that the word was taken in the sense of a geometrical shape. The French and German stimuli were chosen to conform to this sense, but the Italian translation was 'piazza' (a public square), thus making it certain that the Italian primary response could not agree with those of the other languages.

Distribution of agreements. Are these cases of agreement distributed randomly through the list, or is there any regularity in their distribution? A generalization has already been suggested by comparison of French and English primary responses.

> If we take the 25 associations with the highest frequency in French, 22 are equivalent to the corresponding American association. In successive sets of 25 French associations by declining frequency, there were 12, 10, and 4 agreements with the American associations. . . . Thus, the most frequently given associations are shared by the two languages, while the less frequent associations tend to differ.[15]

A similar comparison was made among the four languages. Using the 89 items for which the Italian response is known, the primary responses of a language were divided into halves according to the relative frequency with which they were given. The responses were also categorized according to whether they agreed with zero, one, two, or three corresponding responses of the other languages. The results of this analysis are given in Table III. Let us examine the distribution of agreements between English and the other three languages. Among the 46 English primary responses at or above the median frequency, only 7 do not agree with any of the corresponding primary responses of the other three languages, 9 agree with one of

[13] Russell and Meseck, *op. cit.*, 201.
[14] Rosenzweig, *op. cit.*, *Année Psychol.*, 57, 1957, 24.
[15] Rosenzweig, *op. cit.*, *Amer. Psychologist*, 14, 1959. The quotation is taken from the text of the paper, dittoed copies of which were distributed at the meeting of the American Psychological Association, 1959.

the corresponding responses, 12 agree with two other responses, and 18 agree with all three responses of the other languages. Among the 43 English primary responses whose frequency is below the median, the trend is just the reverse: 24 do not agree with a single other corresponding response, 7 agree with one response, 9 agree with two, and only 3 agree with all three responses of the other languages. Within each language, the more frequent responses show many more agreements than do the less frequent responses. The relationship between frequency of response and number of

TABLE III

AGREEMENTS WITH PRIMARY RESPONSES OF OTHER LANGUAGES, AMONG THE 89 ITEMS WHERE THE ITALIAN RESPONSE IS KNOWN

Language	Percentage	Number of agreeing primary responses				Σ
		0	1	2	3	
English	≥ 33	7	9	12	18	46
	< 33	24	7	9	3	43
French	≥ 18	12	3	14	18	47
	< 18	21	10	8	3	42
German	≥ 18	11	5	10	17	43
	< 18	21	10	11	4	46
Italian	≥ 17	10	9	7	18	44
	< 17	28	7	7	3	45
Over-all	$\geq$ Mdn.	40	26	43	71	180
	$<$ Mdn.	94	34	35	13	176
	Σ	134	60	78	84	356

agreements is significant at better than the 0.1% level in English, in French, and in Italian, as tested by χ^2; in German the effect is significant at better than the 1% level. The distributions for all four languages were then cumulated to give a total distribution. A response in the upper half of the total frequency-distribution agrees, on the average, with 1.81 other responses; a response in the lower half agrees, on the average, with only 0.81 other responses. Of the responses that agree with all three corresponding responses in the other languages, 84.53% are in the upper half of the frequency-distribution; of responses that agree with no corresponding response, only 29.85% are in the upper half of the frequency distribution.

It can also be shown that the mean percentage of response varies monotonically with the number of agreements. The results of this analysis are displayed in Fig. 1. For each language, the responses that agree with none of the corresponding responses of the other languages have a mean percentage only about half as great as that of responses that agree with all three corresponding responses. There is only one minor exception to the monotonic rise of frequency with agreements: in German, the mean percentage for zero agreements is two more than the percentage found for one agreement. When the primary responses of a language were thus classified according to the number of agreements with primary responses of other languages, about one-fourth of the total variance of response-frequencies was accounted for by the variance between classes. The correlation ratios between response-frequency and number of agreements were

0.54 for English, 0.57 for French, 0.48 for German, and 0.45 for Italian. Each of these correlation ratios was significant at better than the 0.1% level of confidence.

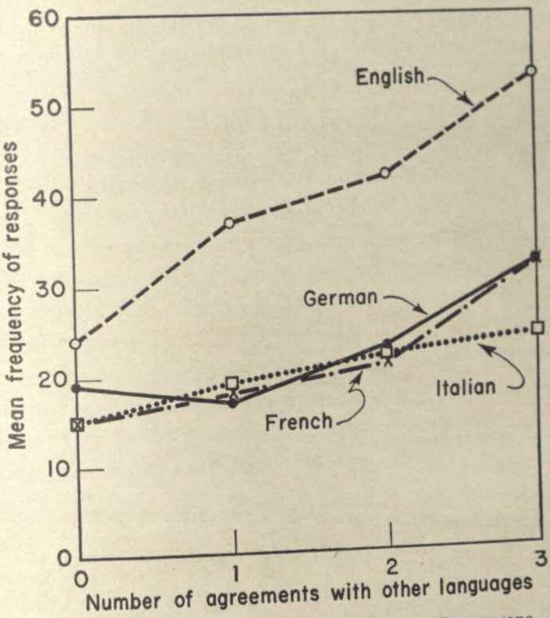


FIG. 1. MEAN FREQUENCIES OF PRIMARY RESPONSES

Fig. 1 also brings out the much greater frequency of primary responses of the American *S*s as compared to the European *S*s. This difference has previously been reported for French[16] and for German.[17] It now appears that all three European groups have rather similar frequencies of primary responses and that the communality of response among American *S*s is almost twice as great.

Response class and agreements. Can classification of the responses according to the relation between stimulus and response aid in predicting agreements among languages? Two frequently used categories were analyzed first, *opposites* (*e.g.* dark-light, sickness-health) and *coördinates* (*e.g.* table-chair, man-woman). Throughout this section we will consider only the 89 items for which the Italian response is known. Among the four languages the *opposite* and *coördinate* classes account for 23 and 37%, respectively, of all primary responses.

[16] Rosenzweig, *op. cit.*, *Amer. Psychologist*, 14, 1959, 363.
[17] Russell and Meseck, *op. cit.*, 200-201.

(1) Opposite responses. The opposite responses were tabulated for each language as in Table III, both according to the number of agreements with the corresponding responses of the other languages and according to whether their frequency was above or below the median. The separate distributions for the four languages were then cumulated, and the results are given in Table IV A. For every language, most of the opposite responses are above the median in frequency; over-all, 77% are above the median. As would be expected of a group of responses with generally high frequency, the opposite responses show a rather high mean number of agreements

TABLE IV

AGREEMENTS WITH PRIMARY RESPONSES OF OTHER LANGUAGES, ACCORDING TO CATEGORIES OF RESPONSES

Freq. response	Number of agreeing primary responses 0	1	2	3		Σ	*M*
A. Opposite responses:							
≥Mdn.	5	3	21	35	64		2.34
<Mdn.	5	2	7	5	19	83	1.63
							2.18
B. Coördinate responses:							
≥Mdn.	12	14	18	27	71		1.85
<Mdn.	31	15	12	1	59	130	0.70
							1.33
C. Adjective-adjective responses							
≥Mdn.	5	5	17	35	62		2.32
<Mdn.	6	5	2	5	18	80	1.33
							2.10
D. Noun-noun responses:							
≥Mdn.	23	17	22	33	95		1.69
<Mdn.	57	21	27	7	112	207	0.86
							1.24

with the corresponding primary responses of the other three languages—2.18, as compared with only 1.31 for all responses taken regardless of classification. Among the opposite responses, the factor of relative frequency of response exerts a significant effect ($\chi^2 = 4.73$, $df. = 1$, $P < 0.05$).[18] The small number of opposite responses below the median frequency makes it necessary to group together the categories 0, 1, and 2 which probably reduces the obtained significance. Those opposite responses that are above the median in frequency show a mean number of 2.34 agreements with the corresponding responses of the other languages. This value is considerably higher than the mean of 1.81 found for all responses above the median, regardless of classification. Those opposite responses with a frequency below the median show a mean number of 1.63 agreements, which is considerably greater than

[18] It may be asked whether the entries in this table are independent, since each response is counted for its agreement or disagreement with each of the three other languages. Responses to the same stimulus are, however, free to fall in different cells of the table. Thus, for Stimulus 29, the French response is not an opposite and so does not appear in this table. The English, German, and Italian responses are opposites and agree with each other. The English response is below the median frequency of English responses and thus appears in the lower half of the table; the German and Italian responses are above the median for their languages, and thus appear in the upper half of the table.

the mean of 0.81 agreements found for all responses below the median in frequency. Thus, the opposites of high and of low frequency show considerably more agreements than do responses of high and low frequency in general.

(2) *Coördinate responses.* Table IV B presents the results for the coördinate responses. Unlike the opposite responses, the coördinate responses show no strong tendency to be high in frequency, only 55% being above the median. Their mean number of agreements with primary responses of other languages is 1.33, compared with 1.31 for unselected responses. Thus, the fact that a response is in the coördinate class does not help in predicting the number of agreements it will have with corresponding primary responses of other languages. Among the coördinate responses, the factor of relative frequency of response exerts a highly significant effect $\chi^2 = 32.93$, $df. = 3$, $P < 0.001$).

Inspection of the opposite responses shows that among the 83 opposites, for all four languages, 73 are adjective-adjective pairs, *i.e.* both stimulus- and response-words are adjectives. Names of colors are considered to be adjectives here. These 73 responses in turn account for almost all of the 80 adjective-adjective pairs found among the primary responses of the four languages. Thus, almost the same result is obtained whether one classifies according to the relational category of opposition or according to the grammatical category, adjective-adjective (A-A).

(3) *Adjective–adjective pairs.* The tabulation of A-A responses by frequency and by number of agreements with responses of other languages (Table IV C) is very similar to that of opposite responses (Table IV A). A χ^2-test shows that the effect of frequency on number of agreements is significant for A-A responses ($\chi^2 = 13.02$, $df. = 2$, $P < 0.01$). The A-A responses, as a group, show a rather high mean number of agreements with other languages, 2.10. This number is very similar to the mean of 2.18 found for opposite responses and is well above the mean for all responses, taken without regard to category.

While the category of opposite responses tends largely to be composed of A-A responses, there is an equally strong tendency for the category of coördinate responses to be composed of noun-noun (N-N) pairs.

(4) *Noun–noun pairs.* Of the 130 coördinate responses, 114 are N-N pairs. Only five coördinate responses are A-A pairs, and these are all pairs of colors. In turn, the 114 N-N coördinate responses make up more than half of the total of 207 N-N responses. Tabulation of the N-N responses by frequency and by number of agreements (Table IV D) gives a distribution somewhat similar to that for coördinates (Table IV B). As was true for coördinate responses, the more frequently given N-N responses have significantly more agreements with the responses of other languages than do the less frequent N-N responses ($\chi^2 = 31.02$, $df. = 3$, $P < 0.001$). While the N-N responses of above-median frequency show a mean of 1.69 agreements with responses of other languages, those of below-median frequency show a mean of only 0.86 agreements. The N-N responses, as a group, have a mean of 1.24 agreements with the responses of other languages.

This is similar to the mean for coördinate responses and is slightly below the mean for all responses, without regard to category.

Discussion

In what terms can we attempt to account for the relation between agreements of primary responses among languages, on the one hand, and frequencies or categories of responses, on the other? Let us consider first the role of frequency, employing the following model: Suppose two random samples are taken from the word-association responses made to a single stimulus by the members of a language community. If the population gives one particular response with far greater frequency than any other response, then both samples will tend to show that response as their primary one. If, on the other hand, the primary and secondary responses of the population are not very different in frequency, then the two samples may well show different primary responses. If the statistics of the population are not known, one can nevertheless predict from one sample to another. The more the frequency of the primary response for an item exceeds the frequency of the secondary response in one sample, the more likely it is that the same response will be primary in another sample. This model predicts quite well the occurrence of the agreements between primary responses of women and men on the French test of word-association. If we take the 25 primary responses with the highest frequencies among the women, all 25 agree with the corresponding primary responses of the men; of the next 25, 23 agree; of the third set of 25 responses, 15 agree; of the 25 responses of lowest frequency among the women, only 12 agree with the primary responses of the men. When a similar analysis is made for the English data of Tresselt, Leeds and Mayzner,[19] the respective numbers of agreements are 25, 23, 21, and 13. Thus, the more frequent the primary response is in one sample, the more stable is it from sample to sample. (The somewhat greater agreement between the sexes in English than in French is consistent with the generally higher frequencies of primary responses in English as compared to French.) While division by sex does not give random samples, there is little evidence of significant difference in the associations of the two sexes, and the results for each language fit the model well. A further test is provided by the distribution of agreements between the English studies of Russell and Jenkins and that of Tresselt, Leeds and Mayzner. Of the 25 most frequent primary responses found by Russell and Jenkins, all 25 agree with the corresponding primary responses found by Tresselt, Leeds and Mayzner. In suc-

[19] Tresselt, Leeds, and Mayzner, *op. cit.*, 150-152.

cessive groups of 25 responses, arranged by decreasing frequency, there are, respectively, 25, 22, and 10 agreements. Again the obtained distribution follows the prediction of the model.

Thus far we have shown that the probability of agreement of primary responses among samples from the same population should be related to the extent to which the primary response exceeds the secondary response in frequency. Only for English have complete norms been published, however, so that the differences in frequencies between primary and secondary responses are readily available only in English. For French, German, and Italian, only the frequencies of primary responses have been published. To see how well the frequency of the primary response could serve as an index of the difference in frequencies between primary and secondary responses, the correlation was obtained between these two values for all 100 items in the English and French norms. The respective coefficients were 0.97 and 0.94, indicating that we are fully justified in using the frequency of the primary response in place of the difference in frequencies between primary and secondary responses.

To apply the model across languages requires the hypothesis that there is a cross-cultural community which shares verbal associations and meanings, even though verbal forms are different. If this hypothesis is limited to western European languages, then it is clearly supported by the results we have obtained. That is, the more frequently a primary response is given in one language, the more stable is the meaning of the response in other languages. It is as if each of the languages we have considered is a sample of a more general language-system.

It is tempting to try to extend the hypothesis to all languages, but data do not yet exist to test adequately this broad form of the hypothesis. A limited and preliminary test can be attempted with a small amount of Navaho material. The material on Navaho word-association was obtained by Ervin and Horowitz in 1956.[20] Their test was composed of 114 items, 42 of which were equivalent to items of the list of Kent and Rosanoff. The test was administered individually and orally in Navaho to 38 adult *Ss*, about two-thirds of whom were women, and most of whom spoke only Navaho. In view of the small size of the Navaho sample and of the difficulties of translation, comparisons with the Navaho material cannot have more than suggestive value. For 3 of the 42 items, the Italian response is not known, so comparisons can be made across all five languages for 39 items. In many cases the Navaho primary response is equivalent to that of the other languages (*e.g.* man-woman, soft-hard, hand-foot, long-short, heavy-light). In other cases, a cultural difference is apparent (*e.g.* bitter-chili pepper, river-canyon, beautiful-horse). The Kent-Rosanoff items included in the Navaho test tend

[20] I am grateful to Drs. S. M. Ervin and A. E. Horowitz for their permission to inspect and cite this unpublished material.

to be predominantly selected from among items showing rather high inter-language agreement among the four European languages. To compare the number of agreements among the primary responses of the five languages, each language was considered in turn and the number of agreements with the responses of the other languages was tabulated for each of the 39 items. The English responses showed a mean number of agreements per item of 2.41; French, 2.05; German, 2.08; Italian, 2.03; Navaho, 1.69. Thus Navaho clearly shows fewer agreements with the European languages than each of the European languages shows with the other four. Nevertheless, the number of agreements between Navaho and any single European language is substantial. It ranges from 21 cases of agreement (out of 39) with English down to 14 cases of agreement with German. This extent of agreement offers support for the hypothesis that speakers of all languages share many verbal associations, even though their verbal forms are different.

In indicating why frequent responses tend to show strong agreement among languages, our model may also help to indicate why opposite or A-A responses show relatively high agreement among languages. If a stimulus has a clear opposite, this response tends to be given with much greater frequency than other responses, as Karwoski and Shachter found: "not all words have opposites; usually there is only one good opposite. When a stimulus-word has a common opposite word, the tendency is for the responses to pile up on that one word."[21] Karwoski and Schachter did not mention grammatical categories, but it appears that adjectives are more likely to have opposites than are nouns. This is certainly the case in dictionary-citations. Thus adjectives are likely to have clear opposites, and if an opposite response exists it tends to be given with considerably greater frequency than other responses. As we have shown, such predominance of one response is the condition in which agreements among primary responses of different languages are to be expected.

If associative habits were completely general among languages, we would expect to find as many agreements of primary responses between two languages as between two samples from the same language. We have seen that this is not the case. Between English and French, for example, there are 48 cases of agreement, while between French men and women there are 75 cases of agreement. Is the entire difference between these figures to be accounted for by differences in English and French associative habits? There are at least two other factors to be considered:

(a) Part of the difference appears to arise from an artificial cause—difficulties of translation. In at least 10–15 cases, the French stimulus-word was not an exact equivalent of the English, and about the same proportion seems to hold for German and Italian. We have mentioned above the difficulties with 'sheep' and

[21] T. F. Karowski and J. Schachter, Psychological studies in semantics: III. Reaction times for similarity and difference, *J. soc. Psychol.*, 28, 1948, 133.

'mutton' in French and with 'square' in Italian. In several other cases, the word translated was not of the same grammatical form as the English word. There is a strong tendency for response-words to be of the same grammatical form as the stimulus-word. Thus, changing the form of the translated word from that of the English makes it unlikely that the responses in the two languages will agree.

(b) In some cases, a difference occurred when the primary response in one language was one of completion, *i.e.* a complex of which the stimulus-word is a part. This is a type of associative habit, but of a special nature that deserves separate mention. It is extremely rare that two languages share the same completion-response, even though their other associations to the stimulus-word are similar. Thus, the English primary response to 'whistle' is 'stop' ('whistle-stop'). French does not have this compound term. The French primary response is 'train,' which is the secondary response in English. Each of the four languages has two or three cases in which the primary response is one of completion. Thus, between a pair of languages, about five disagreements will arise from this cause.

We can then apportion the cases of disagreement of primary responses between a pair of languages as follows: (a) About 25 cases are due to instability of the primary response, even samples from the same language failing to agree in this many cases. (b) In another 10–15 cases, the translation of the stimulus-word is inexact, so the similarity of association between languages is not tested adequately. (c) In about 5 cases, one language shows a completion-response and a similar complex does not exist in the other language. Thus, between a pair of Western European languages, there seem to remain not more than perhaps 10 cases in which cultural factors give rise to a substantial difference in associations.

One strong qualification must be attached to our hypothesis that speakers of different languages tend to share the same associative habits—it holds only among adults. Comparison of responses of American children with those of American adults have shown very different types of responses.[22] It would be interesting to see whether children speaking other languages respond similarly to American children, as we have shown adults of other languages to respond similarly to American adults. Such a study could help in the search for general factors operating in the learning of language in all cultures and tending to produce similar associative habits regardless of the forms of particular languages.

SUMMARY

Comparisons were made among primary responses to the word-list of Kent and Rosanoff and to its translations in French, German, and Italian. There was a strong tendency for primary responses to corresponding stimulus-words to be equivalent in meaning. In each language, the greater

[22] R. S. Woodworth, *Experimental Psychology*, 1938, 344-347.

the frequency with which a particular primary response was given, the more likely was that response to agree in meaning with the corresponding primary responses of the other languages. When both stimulus- and primary response-words were adjectives or when the response was opposite in meaning to the stimulus-word, there also was high agreement among languages. (The two categories, 'opposite' and 'adjective-adjective' had seven-eighths of their cases in common.)

The results can be accounted for on the assumption that, across languages, similar associations tend to occur among words of similar meaning, regardless of differences in verbal forms among the languages. Within a single language, the more a particular response tends to predominate in one sample of responses to a given word, the more likely it is that the same response will be primary in another sample. If associative tendencies are shared among languages, then the more a particular response predominates among the responses to a given stimulus-word in one language, the more likely it is that an equivalent response will be primary in another language.

CHANGES WITH AGE IN THE VERBAL DETERMINANTS OF WORD-ASSOCIATION

By SUSAN M. ERVIN, University of California, Berkeley

The earliest investigators of word-association noted that adult associations were usually in the same grammatical class as the stimulus-words, and that such responses had especially short latencies.[1] These findings challenge an explanation of the learning of word-associations through simple contiguity in overt speech. On this basis, the most frequent response to transitive verbs would be *the*.

A paradigmatic response—that is, a response in the same grammatical class—might arise through similarity of referents, common affixes, or common past verbal contexts.[2] The last determinant has the greatest generality. Even the isolated words offered in an association-test have been encountered before in verbal contexts. Two words may be said to be contextually similar to the degree that their past verbal environments overlap. Contextual similarity thus includes both grammatical and semantic similarity between the stimulus-word and response-word, to the extent that the totality of verbal contexts defines meaning.

Two models could account for the learning of association between contextually similar words. A forward association would predict that repetition of *a cup of coffee* and *a cup of tea* would lead to the association of *coffee* with *tea* and of *tea* with *coffee,* due to their contiguity during competition

* Received for publication March 4, 1960. These data were collected as part of a project on children's reasoning sponsored by the Higgins Fund. The assistance of Edith Kaplan and Roseanne Mandler is gratefully acknowledged.

[1] Gustav Aschaffenburg, Experimentelle Studien über Assoziationen: I. Die Assoziation im normalen Zustande, *Psychol. Arbeit.* 1, 1895, 209-299; B. B. Bourdon, Observatives sur la reconnaissance, la discrimination et l'association, *Rev. Phil. France et l'Etranger,* 40, 1895, 153-185; C. G. Jung, *Studies in Word Association,* 1919, 234-235; Paul Menzerath, Die Bedeutung der sprachlichen Geläufigkeit oder der formalen sprachlichen Beziehung für die Reproduktion, *Z. Psychol.,* 48, 1908, 1-95; Friedrich Schmidt, Experimentelle Untersuchungen zur Assoziationslehre, *Z. Psychol.,* 28, 1902, 65-95; Albert Thumb and Karl Marbe, *Experimentelle Untersuchungen über die psychologishen Grundlagen der sprachlichen Analogiebildung,* 1901, 1-87; Arthur Wreschner, Die Reproduktion und Assoziation von Vorstellungen, *Z. Psychol. Ergbd.,* 3, 1907, 329-599.

[2] The syntagmatic-paradigmatic distinction was made in discussion of word-association by J. J. Jenkins, in C. E. Osgood and T. A. Sebeok (eds.) *Psycholinguistics, Supplement, J. abnorm. soc. Psychol.* 52, 1954, 114-116; Sol Saporta, Linguistic structure as a factor and as a measure in word association, Minnesota Conference on Associative Processes in Verbal Behavior, 1955, 210-213.

of the response. A reverse, or mediated association, would be learned with practice of *front door* and *back door.* In this case, though, since *door* mediates the association of *front* and *back,* it would be the most likely response in free association. Savings in paradigmatic responses would be expected only in a condition of constrained response. It has been shown that both of these conditions do produce learning.[3]

In the following study, differences with age in the frequency of different types of word-association will be examined with reference to an analysis of learning. The following changes with age are expected.

(1) Decrease in syntagmatic responses. Syntagmatic (sequential) associations are more probable where the variety of contexts following the stimulus-word is low relative to its frequency, reducing the number of competing associates. With age, there is an increase in the length and variety of sentences,[4] so that the relative strength of the average syntagmatic association is less.

Syntagmatic associations refer here to any sequential associate, not necessarily the immediately contiguous one. Since determiners (*the, my*), pure prepositions (*of, from*), copulas (*is, become*), nominative pronouns, and coördinate conjunctions virtually never occur in utterances of one word,[5] they would not be expected as responses on an association-test,[6] and may be eliminated in calculation of the sequential probabilities of stimulus- and potential response-words.

(2) Increase in paradigmatic responses. Paradigmatic associates are more likely when the variety of verbal contexts of a stimulus-term is high relative to its frequency. In addition, as vocabulary increases, children have more contextually similar responses in antonyms, synonyms, and words drawn from the abstraction-hierarchy of the stimulus-word.

(3) Differential shift toward paradigmatic responses. If a word occurs frequently in the final position of a sentence, it has relatively weak syntagmatic associations.

---

[3] Evidence for the model of forward association was offered by W. E. Jeffrey and R. J. Kaplan, Semantic generalization with experimentally induced associations, *J. exp. Pyschol.,* 54, 1957, 336-338, and P. M. Kjeldergaard and D. L. Horton, An experimental analysis of associative factors in stimulus equivalence, response equivalence and chaining paradigms, Studies in Verbal Behavior, 1960, No. 3 (University of Minnesota), 21-34. Response-competition is less likely in a speaker than in a listener anticipating speech sequences. David McNeill has ingeniously suggested that children's slower verbal responses may not create the conditions for paradigmatic contiguity through anticipation.

[4] M. C. Templin, Certain language skills in children, *Univ. Minn. Child Welf. Monogr.,* 1957, No. 26, 76-96.

[5] In an unpublished tally of a day's conversational speech transcribed by W. F. Soskin, functions-words were rare, except for question-words, in single-word uttterances.

[6] While many functional-words would not appear in isolation in either a text- or conversational-count, they are sometimes produced by students as isolated responses in school-exercises. If they occur in isolation as stimulus-words in the association-test itself, their subsequent probability as an isolated response may be increased, as has been reported by Davis Howes, On the relation between the probability of a word as an association and in general linguistic usage, *J. abnorm. soc. Psychol.,* 54, 1957, 84.

Such words are nouns, adverbs, adjectives, and intransitive verbs. On the other hand, adverbs of frequency (*always, seldom*), transitive verbs, and question-words (*when, who*) occur less often in final positions. They would be expected to elicit syntagmatic responses at later ages than other categories.

(4) *Decrease in clang-responses.* Since children have less practice than adults in both verbal and non-verbal associations with words, they are more likely to respond to the immediate sound-properties of verbal stimuli, as if the words were nonsense. Indeed, children are known to display more generalization between words that sound alike.[7]

Earlier research partially supports these generalizations. Wreschner, using German words, found that age and education were both related to paradigmatic dominance, and that children preferred concrete nouns as responses regardless of the class of the stimulus-word.[8] Inflectional affixes in German confound clang-responses and paradigmatic responses.

Reanalysis of Woodrow and Lowell's data for English reveals that paradigmatic responses increased and syntagmatic responses decreased with age.[9] The range of form-classes included was limited, however, and the list was not confined to cases where the adult primary response was present in children's vocabularies. Age-changes may be merely due to specific changes in vocabulary in such cases, though it is true that these alone should not produce the bias in the direction of change which the data indicate.

METHOD

Materials. The list of associative words, 46 in number, was chosen from a variety of grammatical classes. Since the work was done in conjunction with a study on learning antonyms, 39 of the items were so chosen that the primary response of adults was coördinate or antonymous. All stimulus-words and primary responses of adults were, according to Rinsland, probably within the vocabulary of the youngest children.[10] The form-classes were alternated within the list and the same order of presentation was used for all *Ss*.

The *closed-alternative test* was composed of 35 items, the last 10 being omitted for the third-graders. Three kinds of items were alternated on the list to test the relative strength of syntagmatic, paradigmatic, and antonymous responses. In one set the words were grammatically alike (*snow, winter, summer*), but contained an antonym. In a second set, an antonym was contrasted with a syntagmatic associate, (*pillow, soft, hard*). In the third group, syntagmatic and paradigmatic associates were contrasted, (*fire, hot, warm*).

Subjects. Twenty-three *Ss* were chosen from the kindergarten, 10 from the first

[7] B. F. Riess, Genetic changes in semantic conditioning, *J. exp. Psychol.*, 36, 1946, 143-152.

[8] Wreschner, *op. cit.*, 70.

[9] Herbert Woodrow and Frances Lowell, Children's association frequency tables, *Psychol. Monogr.*, 22, 1916 (No. 97), 81.

[10] H. D. Rinsland, *A Basic Vocabulary of Elementary School Children*, 1954.

grade, 52 from the third grade, and 99 from the sixth grade. In the last two grades, the entire class was tested in a group; in the first two, the children were tested individually.

Procedure. The *Ss* in the kindergarten and first grade were tested orally; those in the third and sixth grades gave their responses in writing. To keep speed uniform and to help slow readers, *E* read the words aloud to the older children.

Instructions. The following instructions were given for the word-association test.

When you hear a word, sometimes it makes you think of another word. If you heard *cat* you might think of *milk*, or *purr*, or *dog*, or *black*—almost anything. What does *cat* make you think of? What does *eat* make you think of? Anything else?

All single word answers were accepted.

For the closed-alternative form also given, the children were told to say or to draw a line showing with which word the middle word of three seemed to go best. An example was given: "Does *brother* go better with *sister* or with *father*?"

Response-analysis of free associates. Both the stimulus-words and the response-words were classified by a method of defining grammatical class derived from Fries.[11] Test-frames or contexts were established; a word was assigned to a class if it could fit into the frames altered by substitutions from the same grammatical classes as the words in the frame. The same word might fall in several classes—*walk* is both noun and verb, for instance. In addition, coders were instructed to try to judge probable contexts according to children's usage. Thus *people* was not classified as a verb.

In the following list of some of the principal classes used in the analysis, the numbers indicate the agreement scores, among the coders' obtained by computing the probability that an item coded in the given class by one coder would also be coded in that class by another coder. Disagreements largely stemmed from cases where one coder thought a usage too rare to include.

(1) Nouns, including verbs with '*-ing*', if they could be preceded by adjectives but not adverbs; 0.97.

(2) Pronouns, excluding possessives; 0.96.

(3) Transitive verbs; 0.87.

(4) Intransitive verbs; 0.87. Since most transitive verbs also occur in intransitive contexts (*he likes to eat*), coders were instructed not to code as intransitive any verbs which could be made transitive merely by adding an object.

(5) Adjectives, including verbs with *-ing* and *-ed* if they can be preceded by modifiers or by both adjectives and adverbs of manner; 0.92.

(6) Adverbs; 0.90.

(7) Nominal adverbs; 0.79. This is a subgroup of adverbs which can occur after certain prepositions, *e.g. now, here.*

The above comprise Fries' parts of speech. Below are several classes, of functional words most of which cannot occur alone in an utterance without the presence of some other parts of speech.

(8) Modifiers, which precede adverbs or adjectives; 0.85. Examples are *quite, really, too, just, very.*

(9) Determiners, which include the traditional articles and possessive pronouns; 0.97. Examples: *most, the, my, that, some.*

---

[11] C. C. Fries, *The Structure of English,* 1952, 65-109.

(10) Prepositions; 0.99.

(11) Interrogative words and *subordinate conjunctions* which were pooled for this analysis because of the large overlap in composition of the classes; 0.87.

Sequential analysis. As an approximation to the sequences of grammatical classes in children's speech, children's books were subjected to an analysis of grammatical sequences. One hundred or more cases of items in each class were tallied in sequence from the texts to yield an estimate of the probability of each class given another class, and the probability of occurrence of the class in final position. Tallies were discontinued when it appeared that the probabilities for the given class were stable. Two tallies were noted for each item, one of the class of the word immediately following in the text, and the other of the class of the next word omitting functional words. The only major deviation from Fries' categories was that copulas (*is*, *seems*) were treated as functional words.

RESULTS

Closed-alternative test. Of the 9 items where there were no antonyms so that only form-class was at issue, there were 5 showing significant increase with age in the selection of the paradigmatic alternative (Table I). On the 11 items where antonyms were compared with syntagmatic choices, there were 5 significant increases with age. In all, two cases of marked de-

TABLE I

PROPORTIONAL CHOOSING SAME GRAMMATICAL CLASS ON CLOSED ALTERNATIVE TEST

Stimulus-sets (with correct pair marked)	Kindergarten and first grade ($N=33$)	Third grade ($N=95$)	Sixth grade ($N=56$)
fire hot —warm	61	30	48†
witch wicked—bad	42	35	56‡
up —high sky	39	42	56†
lie —cheat bad	36	47	72‡
black —dark night	27	40	48
trees —grass green	17	—*	18
ball —bat play	66	—	96§
supper eating —drinking	37	—	36
write desk —table	83	—	73
pillow soft —hard	21	15	32†
sad —happy fun	15	37‡	52†
dark night —day	39	39	45
fast run —walk	36	26	34
go —come here	42	58	70**
behind back —front	39	47	48
to —from away	27	41	54**
he him —her	70	42	48
light float —sink	67	63	66
softer —harder stone	24	—	50†
played worked hard	34	—	50

* Some items were omitted at the end of the test in the third grade.
† Higher proportion than in next youngest group, $p<0.05$.
‡ $p<0.01$. § $p<0.001$.
** Higher than youngest group, $p<0.01$, but not different from middle group.

crease appeared, both involving the smallest sample, where the proportions are least reliable.

Paradigmatic responses in free-association. In counting paradigmatic associates on the free-association test, coders used a simple criterion to isolate the purest instances of paradigmatic responses. Responses were called paradigmatic only (a) if they can occur in the same class as the stimulus-word even if each also occurs in other classes, and (b) if they do not occur in immedate sequence or separated only by a determiner in ordinary continuous speech. Thus, though all are nouns, *front-door* and *table-spoon* were not tallied as paradigmatic, nor was *game-play*. The second restriction was not generalized to all function-words rather than merely determiners, because such a rule would include conjunctions. Almost any pair of words in the same grammatical class might, of course, occur in sequence linked by a conjunction. The restrictive rule was pragmatic in origin and in practice successfully excluded the cases which appeared ambiguous as to paradigmatic or syntagmatic status.

One-tailed tests of significance were made of the increases, and are presented only for adjacent groups except when there was a gradual increase that was not sharply inflected. Of the eight nouns, six showed significant increases between the youngest *S*s and those in the third grade (Table II). The remaining two stimulus-words, *moon* and *winter,* had a very high proportion of paradigmatic responses in the youngest group. Five of six verbs, and all of the prepositions, adverbs, pronouns, and comparative adjectives showed increases with age. *When* produced a low proportion of paradigmatic responses at all ages.

The primary responses of adults for most of the stimulus-words was a coördinate or antonymous response. It could be argued that the increase in paradigmatic responses simply represents a learning of a particular type of paradigmatic response, or that it merely represents a culturally stereotyped learning of the primary response or training with conjunctive phrases. There are two tests of this explanation. One consists of examining the words which do not have high-frequency primary responses in the oldest group; namely, *game, build, across, when*. Three of the four showed age-increases.

A more stringent test consists of examining only responses other than those that are primarily adult, thus reducing the sample size in each grade. Fourteen items remain with sufficient cases for a statistical analysis, and of these eight showed significant increases with age in paradigmatic responses.

Thus the shift to paradigmatic responses cannot be regarded simply as

due to increased learning of the adult primary and peculiar to coördinate or contrast responses.

Since the number of multiple-word responses decreased with age, it could be argued that the change between the youngest *Ss* and those in the third grade is largely a result of learning to isolate single words. The

TABLE II

PROPORTION OF PARADIGMATIC RESPONSES ON FREE-ASSOCIATION TEST

Stimulus-word	Response-class tallied	Kindergarten and first grade ($N=33$)	Third grade ($N=98$)	Sixth grade ($N=52$)
table	N	18	69§	79
moon	N	69	80	71
boy	N	69	92‡	98
front	N	57	85‡	77
night	N	57	74*	83
winter	N	69	77	83
hand	N, TV	48	74†	79
game	N	15	33*	34
build	TV	15	18	48§
give	TV	15	34*	63‡
float	TV, IV	33	46	81§
worked	TV, IV	39	55	58
come	IV	33	71§	75
walking	TV, IV	30	52*	81‡
from	P	12	15	44§
across	P, Adv	12	41†	54
over	P, Adv	36	60†	65
up	P, Adv	63	81*	77
out	P, Adv	42	78§	77
before	P, Adv, SC	33	78§	85
always	Adv	12	49§	75†
there	Adv	18	67§	58
yesterday	N, Adv	63	90‡	94
him	Pro	30	76§	71
these	Pro, D	18	64§	71
when	Q	12	14	21
softer	A	42	69†	73
hotter	A	51	71*	62
slower	A	57	79†	81
worse	A, Adv	42	86§	87

* Higher proportion than in next youngest group, $p<0.05$.
† $p<0.01$. ‡ $p<0.001$. § $p<0.0001$.

phrase-responses were retained in the tallies on the grounds that these were the younger child's version of a syntagmatic response, most often consisting of the stimulus-word embedded in a phrase. For the older children the response was often merely the main, or modified, term of the phrase. Thus, in young children *across* might elicit *across the street;* in the third-graders simply *street.*

If only the single-word responses be tallied as a stringent test, a few of

the age-changes disappear and all are of course reduced in magnitude. It remains substantially true even of the single-word responses that paradigmatic frequencies increase.

Syntagmatic responses. The fact that paradigmatic responses increase with age does not demonstrate that the remaining responses bear a systematic relation to the stimulus-words. They might be randomly selected from all grammatical classes regardless of the stimulus-class.

If the response-words in association were a function of immediate succession in texts, 45% of associations to transitive verbs would be determiners, and 51% of associations to intransitive verbs would be prepositions. The actual associative probabilities are, in each case, 2%.

TABLE III

PROBABILITY OF GRAMMATICAL CLASSES IN TEXTS AND WORD-ASSOCIATION FOLLOWING SPECIFIED ANTECEDENT CLASSES*

Response-class	Noun, Pronoun		Intrans. Verb		Trans. Verb		Adjective	
	text	word assoc.	text	word assoc.	text	word assoc.	text	word assoc.
Noun, pronoun	.31	(.70)	.68	.35	.92	.75	.79	.56
Intransitive verb	.19	.21	.01	(.46)	.03	.02†	.01	.06
Transitive verb	.33	.27	.05	.09†	.01	(.46)	—	.13
Adjective	.13	.17	.06	.18	.02	.02	.16	(.71)
Adverb	.03	.02	.20	.28	.02	.07	.03	.03

* Word-association probabilities include function-words and therefore do not add to unity. Parenthesized values represent uncorrected responses in paradigmatic classes.

† When a verb-response to a verb-stimulus was double-coded as both transitive and intransitive, only the paradigmatic code was counted, thus decreasing the values in these cells. Other double coding was not adjusted.

When functional words are omitted from the textual count, on the grounds that functional words rarely occur in isolation and will, therefore, be improbable as associative responses, then the text-sequential probabilities are as shown in Table III. The associative responses shown below them correspond closely, when corrected so that the probabilities of responses in the same form-class as the stimulus-word are based on the expectations if the responses are syntagmatic.

Since there is no way of knowing from inspection which of these responses were based on substitution and which on sequence, an arbitrary correction was made. The probabilities of same-form-class associative responses were assumed to equal those in the texts.

This table is distorted by two factors. One is not easily modified—double coding was used for many responses, and influenced particularly strongly the classes of nouns and verbs. The second difficulty is that there is a difference in the basal probability of each form-class, regardless of

antecedents, in texts and in association. This is to be expected because of the difference between textual frequency and occurrence in single-word utterances. It may be noted that the difference was greatest for nouns and pronouns, chiefly because the latter were much less common in associations than in texts.[12] A correction may be made by transforming the table into deviations from row-means, omitting the diagonal cells. Table III has been presented here in uncorrected form because the correction could be made if desired, yet the raw form may be more useful to other investigators. The product-moment correlation between contingent probabilities in texts and in associations after these corrections was 0.87. This correlation may be interpreted as a measure of the dependence of associations on the class of the stimulus-word, once corrections for paradigmatic probabilities and for differential probabilities of offering a response in text and in isolation have been applied. Thus, most of the variance remaining in the word-associations was due to training in textual sequences.

A few examples may clarify the character of syntagmatic responses. The most common response of changed form-class to *across* was *street,* to *float* was *boat,* to *come* was *here,* to *build* was *house,* to *game* was *play,* to *table* was *eat,* to *when* was *now.* The last may be regarded as syntagmatic in the sense that it represents a response to a question.

Some of these responses seem to indicate backward associations, but the question of associational direction cannot be solved readily with these data, and backward associations undoubtedly confound the data of Table III.

There were very similar distributions of probabilities in the three age-groups when contingent associations were separately examined. There was a significant change with age in the direction of increasing frequencies of transitive verbs as responses to nouns and decreasing adjectival responses to nouns. Possibly this change reflects a change in speech away from descriptive sentences.

Sentence-final position. It was expected that words which can occupy the final position in a sentence may have less strong subsequent syntagmatic responses. In terms of the proportion in final position in texts, adverbs were highest (0.36), next nouns (0.20), and intransitive verbs and adjectives (0.14). The probability of occurrence in final position of transitive verbs and function-words was less than 0.03. On these grounds we would expect that *build, give, from,* and *when* would have strong syntag-

[12] In a day's conversational transcript kindly supplied by W. F. Soskin, pronouns occurred in single-word utterances proportionally less often than any other class except functional words. In proportion to their total frequency, nouns were 7.6 times as probable.

matic associations, and thus less paradigmatic dominance. In addition, *always,* while classed as an adverb, occurs most typically in a pre-verbal position or before an adjectival predicate. It may be seen in Table II that these words all showed late development of paradigmatic dominance, there being a significant increase between the third and sixth grade in paradigmatic responses to *build, give, from,* and *always.* The proportion of paradigmatic responses to *from* and *when* remained low even in the sixth grade. The proportion of paradigmatic responses in the youngest group was highest for nouns, and next for adjectives.

Clang-associates. Clang-associations are interpreted most broadly as all responses with the same initial consonants, with rhyming vowels, or with similar syllables included. Two exclusions were made from the tally—stimulus-words which had clang-antonyms or inflectional affixes such as *-ing, -er,* and *-ed.* With these restrictions, the average number of clang-associates per child decreased from 8.33 in the youngest group, to 2.73 in the third grade and 1.62 in the sixth grade. When only clang-responses which were nonsense words or bore no meaningful link with the stimulus-word were considered, the difference was more marked, the frequencies being 4.39, 0.36, and none, respectively.

Discussion

The marked increase in paradigmatic responses with age might be a result of several factors other than the relative strength of conflicting syntagmatic responses. Older children may have more practice in single-word responses than the youngest. The youngest were tested by an oral technique while the older groups wrote their responses. Yet in the controlled-choice test, all 14 sets showed increases in written paradigmatic responses between the third and sixth grades.

Perhaps the change as a function of age reflects educational experience. First-graders in the schools sampled were using exercise-books practicing substitutions of antonyms and synonyms in sentences. Such exercises were not in use at the time of the study of Woodrow and Lowell, in which the children 9–12 yr. of age showed as high syntagmatic predominance as the kindergarteners in the current study. The widespread use of such materials in this country might account both for the increase in common (usually paradigmatic) responses found by Jenkins and Russell over a 30-yr. period in college students,[13] and for the lower degree of commonality of response

[13] W. A. Russell and J. J. Jenkins, The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff test, Technical Report 11, University of Minnesota, 1954.

in Europeans.[14] Yet this explanation does not suffice completely. Woodrow and Lowell's adult sample did show many paradigmatic responses, and non-literate Navaho adults also markedly prefer paradigmatic primaries.[15] Formal educational practices merely hasten changes which occur with experience even without schooling.

Associational direction. The analysis by classes clearly supports the assumption of predominance of forward associations. There was no evidence, for example, of an increment to noun-adjective or to intransitive verb-noun associative responses arising from backward association, when the textual probabilities are compared to the associative responses.[16]

There is some evidence also for this directional bias in English usage. If a word rarely occurs in the final position of a sentence, it also rarely occurs alone. These classes—nominative pronouns, copulas, and function-words—seem to be the most structurally dependent, apparently because some syntagmatic association is very strong. Thus nouns, intransitive verbs, adjectives, adverbs, and accusative pronouns, which can occur in the final position of a sentence, also appear alone in answers to questions. When, however, questions seem to demand a response in a structurally dependent class, two kinds of answers are given. The answer may be longer than a single word: *who's coming? I am. We are.* Or the respondent chooses a class that can occur alone: *Who's coming? me; whose is it? mine. I, we,* or *my* did not occur alone as a response in spoken texts. The only exception to this pattern is the interrogative-word.

A full test of the hypothesis that frequency of final position is related to an earlier increase in paradigmatic responses should be conducted with stimulus-words which can occur in isolation. Verbs with known positional probabilities would be good candidates for such a study, as would subclasses of adverbs. Adverbs of frequency of occurrence typically precede the verb whereas adverbs of place, manner and absolute time more often occupy sentence-final position.

Of the two models presented earlier, the results of the present study conform to the learning model for forward associations. Can this model also account for other forms of verbal behavior showing similar age-changes? The norms of the Stanford-Binet test and Werner and Kaplan's study of

---

[14] M. R. Rosenzweig, Comparisons among word-association responses in English, French, German, and Italian, this JOURNAL, 74, 1961, 347-360.

[15] Sample of 38 Navahos collected for the Southwest Project in Comparative Psycholinguistics by Arnold Horowitz and Susan Ervin.

[16] If the data in Thumb and Marbe, *op. cit.*, 56-63, are retabulated separately for intransitive and transitive verbs, verb-responses dominate for the former, and nouns for the latter.

nonsense-words both showed age-changes in definitions.[17] The younger children offered sentences as definitions; the older offered synonyms. Brown and Berko have shown that there is a high correlation between paradigmatic dominance in associations and synonymous definitions, for various grammatical classes.[18] The simple model of forward-association does not seem adequate to account for synonymous definitions of nonsense-words such as those found by Werner and Kaplan, but experimental evidence is not available on this point.

SUMMARY

Children in kindergarten, first, third, and sixth grades were given free- and two-choice associative-tests. It was found that there was a significant increase with age in the proportion of responses in the same grammatical class as the stimulus-word, with an earlier increase in words occurring more often in final position in sentences than in words typically medial in sentences. There was a decrease with age in clang-associations. When paradigmatic associates were removed, there was a correlation of 0.87 between the transitional probabilities of five grammatical classes in word-association, and the five classes in texts with the function-words or connective words eliminated. The functional-words do not ordinarily occur in isolation in speech and virtually never occurred as response-words in association.

These findings support a theory of associations based on training by forward contiguity in speech. Responses in the same grammatical class as the stimulus can be learned on the basis of occurrence in the same preceding verbal contexts. Their predominance over sequential associations from speech could derive from the relative variety of the contexts of the stimulus-word, and from the relative strength of substitutable terms. Both contextual variety and size of vocabulary increase with age and hence responses should come increasingly to correspond in grammatical class to the stimulus-word.

[17] Heinz Werner and Edith Kaplan, The acquisition of word meanings: A developmental study, *Soc. Res. Child Developm. Monogr.*, 15, 1952 (No. 51), 84.

[18] R. W. Brown and Jean Berko, Word association and the acquisition of grammar, *Child Development*, 31, 1960, 1-14. This study also replicated the findings on change with age reported here.

OBJECTIVE MEDIATORS IN PAIRED-ASSOCIATE LEARNING

By ALLAN BARCLAY, Saint Louis University

The various hypotheses which have arisen with respect to mediational variables in the modern theories of learning, although differing in details, are quite similar. They derive historically from the classical associationists' position on the problem of mediate association. Stated in its simplest form, this position holds that an associative connection between two terms, not themselves directly associated, is mediated by a third term which bears an associative relation to the original terms.[1] This position allowed the associationists to deal more readily with complex cognitive processes. For example, it allowed them to account for the occurrence of 'new' ideas, which seemed to occur in the absence of any known associative relation between the 'new' idea and its immediate predecessor, on the basis of some actual association which the 'new' idea and its predecessor shared in common. This common association (or in more modern terms, common response) is the basis for a mediational explanation which is widely current in experimental work on learning.[2]

For the most part these studies accept, in one form or another, and often implicitly, the interpretation of mediational effects in terms of common associations. Although this explanation has been accepted by many experimenters, the actual operation of these intermediaries is unknown. The problem of how mediation occurs in association is crucial. It has been difficult, however, since the presumed mediators are considered unobservable,[3] to formulate an experimental design in which the effect of mediation *qua* mediation could be discovered. This lack of observable mediators has led to the proliferation of speculative hypotheses about their nature and operation.

The present study addressed itself to an evaluation of mediation in

* Received for publication May 31, 1960. This article is based on a dissertation submitted in partial fulfillment of the Doctor of Philosophy degree in the Department of Psychology of Washington University. Directed by Dr. J. M. Vanderplas.

[1] William Hamilton, *Lectures on Metaphysics*, 1859, Lectures XVIII and XXXII; L. A. Selby-Bigge (ed.), *David Hume: A Treatise of Human Nature*, 1888, Section IV.

[2] See C. L. Hull, Knowledge and purpose as habit mechanisms, *Psychol. Rev.*, 37, 1930, 511-525, and C. E. Osgood, Theories of learning, *Method and Theory in Experimental Psychology*, 1953, 392-412.

[3] Osgood, *op. cit.*, 411.

paired associative learning to determine whether effects often attributed to mediation are indeed due to mediation *per se* or are attributable to other sources within the framework of the experiment.

METHOD

Experimental design. The experimental design for this study was essentially a mediate associational design with provision for the inclusion of an objective mediator. The paradigm is indicated in Table I, where *Sb* and *Sa* are random shapes and *Rc* and *Rd* are nonsense-syllables, both having known associative values.[4] The contingency of the objective mediator *Sb* in the experimental groups was dependent upon *S*'s correct response to *Rd* and was presented on a random schedule according to which level of associative strength of the mediator obtained for the given experimental group. The experimental method was the traditional one of learning by paired-associate, with the exception that in order to render the mediator objective, if *S* responded with a correct *Rd*, the random shape serving as the mediator (contingent *Sb* in the experimental paradigm) was presented contiguously with the correct response *Rd*, *i.e.* the *Sb* in Task 2 trials was contingent upon a correct response and was under experimental control. Thus, presumably, on Task 3 trials the following state of affairs existed; *Sa* — (*Rd* · · *Sb*) — *Rc* where *Sa* evoked *Rd* because of the previously learned associative connection and in turn, *Rd* evoked *Sb* by virtue of their contiguous associative pairing. Since *Sb* was associated with *Rc* in the prior experience of the *S*, *Sb* should have evoked the response *Rc* with greater facility than if such an objective mediator were not available, as in the control group. Note that in Task 3 trials the stimulus-member *Sa* should tend to evoke the sequence leading to the correct *Rc* response not because of an inferred associative connection but because of known prior associative connections. Thus, the mediator became objective and subject to experimental control.

TABLE I

EXPERIMENTAL PARADIGM

Experimental		Control	
Task 1: *Sb*—*Rc* Task 2: *Sa*—*Rd* . . . *Sb* Task 3: *Sa*—*Rc*	Four experimental groups; contingent *Sb* present on either 25%, 50%, 75%, or 100% of correct *Rd* responses	Task 1: *Sb*—*Rc* Task 2: *Sa*—*Rd* Task 3: *Sa*—*Rc*	Contingent *Sb* not presented

Subjects. Fifty-five university students, randomly assigned to the experimental and control groups, were used as *S*s. All were volunteers from summer school classes and were equivalent in age and educational status.

Materials and apparatus. Stimulus-materials for the study consisted of eight-point random shapes with an average associative value of 37% and were printed

[4] The identification of stimulus- and response-terms by S and R, as well as the identification of the materials to be learned by lower case letters, a, b, c . . . x, will be used consistently throughout this paper.

photographically on high contrast paper (black on white).[5] Response-materials consisted of nonsense-syllables, with an average associative value of 36%, constructed into lists according to standard rules.[6] The experimental apparatus consisted of a Lafayette memory drum modified to permit contingent presentation of the mediating stimulus.

Procedure. The method of paired-associates was used throughout. Each stimulus-member, exposed for 2 sec., was followed by a 2-sec. exposure of the stimulus- and response-members together. A 4-sec. interval elapsed between successive trials. Learning was carried to a criterion of three consecutive errorless recitations.

Instructions. All the *S*s were instructed as follows:

This is an experiment on learning. A series of nonsense-shapes and nonsense-syllables will be shown to you through this window. A nonsense-shape is a shape without any meaning just as a nonsense-syllable is a combination of letters without any meaning. I wish you to look at these shapes and syllables carefully as they appear in the window. After you see the shape alone for 2 sec., the drum will turn and show the shape together with a syllable for another 2 sec. There are a number of such pairs in this list which will be shown to you in this manner. Your task is to learn what syllable goes with what shape and to spell out for me that syllable, before the drum turns. The first time through the list I wish you to learn as many of the pairs of shapes and syllables as you can. After the first trial, the same list appears over here but in a different order. Remember, your task is to spell out each syllable before it appears in the window. Don't be afraid to go ahead and spell out what you think the syllable is even if you're not sure that it is right. It will be easier for you to learn this list if you begin trying to spell out the syllable right away. It will take several trials before you learn the list perfectly but do not get discouraged. Keep trying and do as well as you can. Ready?

For the experimental groups the following additional instructions were given at the time of presenting the second list, on which the contingent mediator would appear:

On this next list from time to time, as you give a correct response, another nonsense-shape will appear over here next to the syllable. I wish you to pay attention to this nonsense-shape, but do not let it distract you from learning this list. Try to remember it as well as you can. It will not appear every time but it will appear often and I wish you to notice it carefully when it does appear.

The method of presentation was, in some measure, the crux of the experiment since it was necessary to present the contingent mediator contiguously with a correct response to provide for an objective and observable mediator. To present this contingent mediator contiguously with the correct response, a method was devised wherein the mediator would always be present but only observable to *S* in the event of his correct response, and then only if the random presentation schedule for the level of associative strength of the mediator for a particular experimental group indicated that the mediator was to be presented. The mediator was so placed on the lists of materials to be learned that it always occurred in the paired-associate presentation as a third member of the stimulus-response pair, *i.e.* the mediator was an integral member of the paired-

[5] J. M. Vanderplas and E. A. Garvin, The association value of random shapes, *J. exp. Psychol.*, 57, 1959, 147-154.

[6] E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 517-567.

associate group. To control the presentation of the mediator to *S* the memory-drum was modified by the addition of a shutter so constructed that it would normally show only the regular response-members but when activated by *E* it would show the response-member contiguously with the mediator. This shutter consisted of a flat metal plate, with apertures corresponding to the size of stimulus- and response-members in one position and with apertures corresponding to the size of the response- and mediator-members in its other position, held by a metal track allowing it to move freely in a vertical plane. Thus, in the normal position both

TABLE II

MEANS AND STANDARD DEVIATIONS FOR EXPERIMENTAL AND CONTROL GROUPS, LISTS 1, 2, AND 3

Group	List 1		List 2		List 3	
	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*
E-25	30.09	6.78	17.45	6.56	14.00	4.78
E-50	26.27	8.88	12.27	7.69	13.09	3.88
E-75	27.27	8.97	15.00	5.37	14.00	4.00
E-100	31.18	6.72	16.73	6.62	14.54	3.46
Control	37.54	4.16	24.09	9.49	23.36	7.61

stimulus- and response-members could be seen through the aperture but the mediator was concealed. In the other position the response-member and the mediator could be seen through the aperture but the stimulus-member was concealed. This arrangement allowed the mediator to be presented contiguously with the correct response upon which it was dependent. It was thought that this contiguous pairing was essential to the development of mediation since it provided for an associative connection to be established experimentally which could then serve to mediate faster learning among the experimental groups in subsequent test-trials. Thus, although the essential method was that of traditional paired-associate learning the provision for presentation of the contingent mediator by means of the modifications discussed above offered a method for experimental control of the mediating stimulus-object. It will be noted that the experimental paradigm for this study can be thought of as a traditional mediate associative design but with an extension of the response-term in second-task learning trials. In other words, the whole term Rd . . . Sb which was presented earlier in the discussion of the experimental paradigm is really synonymous with the usual second-task term, but has been so modified as to provide for experimental control of the mediator. This offers a means of testing experimentally certain hitherto untested assumptions as to the nature of mediation and the manner in which it proceeds, and, it further provides for the comparison of mediation and transfer as separable effects. That is, it offers a sensitive method for allowing mediation to develop in the experimental groups as a result of the contingent mediator and by reference to the level of transfer existing in the control group the facilitative effect of the contingent mediator, if any, can be observed.

RESULTS

Statistical analyses. The means and standard deviations of Lists 1, 2, and 3 are presented in Table II. It will be noted that while the means for

the investigated groups are consistent among themselves on all lists, the means for the control group are higher on all lists than any of the means of the experimental groups.

As a first step in the statistical treatment of the data, an analysis of variance of the scores of Task 3 learning was performed. This analysis is presented in Table III.

The *F* obtained indicated that a significant effect existed among the

TABLE III

ANALYSIS OF VARIANCE OF SCORES FOR LIST 3

Source	*df.*	Mean square	*F*
Between	4	199.65	8.04*
Within	50	24.84	
Total	54		

* $p<0.01$.

TABLE IV

ANALYSIS OF VARIANCE OF SCORES FOR LISTS 1 AND 2

Source	List 1			List 2		
	df.	Mean square	*F*	*df.*	Mean square	*F*
Between	4	211.33	3.99*	4	216.02	4.03*
Within	50	53.00		50	53.59	
Total	54			54		

* $p<0.01$.

groups as a result of the experimental procedure. A test of the homogeneity of variance of scores for List 3 was made since the *F* obtained, which was significant, might have arisen as a result of differences among the variances rather than a result of differences among the means.[7] The obtained Chi-square (6.60, *df.* 4) was not significant and indicated that the *F* obtained did not arise from lack of homogeneity among the variances. Inasmuch as the mean score (trials to criterion) for the control group on Lists 1 and 2 was higher, however, than any of the mean scores for the experimental groups, an analysis of variance for these lists was undertaken to determine whether significant differences existed for those lists. The results of these analyses are presented in Table IV.

The *F*s obtained for Lists 1 and 2 indicated that a significant effect did exist among the groups on both lists. This result was unexpected, since random assignment of *S*s to the experimental groups should have produced relatively homogeneous performance on these lists. Such an observed effect could have arisen as a result of differences in variances among groups as well as of differences in means. A test was, therefore, made to determine whether the groups were random samples

[7] A. L. Edwards, *Experimental Design in Psychological Research*, 1950, 195-197.

having a common variance. Tests for homogeneity of variance indicated that the hypothesis of random sampling from a population having a common variance was tenable since none of the observed Chi-squares was significant (List 1: 3.70, *df.* 4; List 2: 8.89, *df.* 4). It may be inferred on the basis of this test that the differences which exist among groups are due to differences among the means and are not due to differences among the variances.

TABLE V

Analysis of Covariance of Scores Controlled for Performance

Source	List 1			List 2			Lists 1 and 2		
	df.	Mean square	*F*	*df.*	Mean square	*F*	*df.*	Mean square	*F*
Total	53			53			53		
Within	49	20.83		49	15.24		49	18.03	
Adjusted	4	86.59	4.16*	4	101.70	6.67*	4	63.01	3.49†

* $p<0.01$. † $p<0.05$.

The finding that significant effects existed both on List 1 and 2 as well as on List 3 meant that it was necessary to develop a statistical control for the differential performance among the groups on the initial lists. That is, the effect of the greater number of trials to criterion for the control group could have been the basis for the significant result found on List 3. If this were true, it was then necessary to ascertain whether the significant result still remained when the effect attributable to differential initial performance-level was statistically controlled. The appropriate technique for this control was the analysis of covariance, and the results of those analyses are presented in Table V.

The results of these analyses indicated that a significant treatment-effect on List 3 remained when controls were exercised separately over first and second list performance as well as when cotnrols were exercised over first and second list performance combined. Thus, the significant effect found for List 3 is not attributable to differences among groups with respect to their earlier performance on Lists 1 and 2. To ascertain the actual mean trials to criterion, for a more sensitive comparison of means when differences among the groups are taken into account, it is possible, however, to adjust the means statistically. This is done by introducing an adjustment which depends upon the regression coefficient within groups and the deviation of the observed mean for a particular group from the grand mean of the total group. The results of this adjustment for the present data are presented in Table VI.

It can be observed that the adjustment of the means of the combined experimental groups and the control group, to allow for the disparity

originally existing with respect to mean scores, still reveals a difference in trials to criterion. To evaluate these mean differences, *t*-tests were made. The results indicated that a statistically significant difference existed between the adjusted means of the combined experimental groups and the adjusted mean of the control group (List 1, $t = 6.55$, *df*. 55, $p < 0.01$; List 2, $t = 6.51$, *df*. 55, $p < 0.01$; List 3, $t = 10.27$, *df*. 55, $p < 0.01$). A further inspection of these adjusted means indicated that order-effects attributable to the strength (probability of occurrence) of the mediating association were not present. Additional *t*-tests comparing the weakly associated mediation group (E-25) with the strongly associated mediation

TABLE VI

ADJUSTED MEANS, CONTROLLED FOR LISTS 1 AND 2 COMBINED

Group	Observed means		Regression coefficient	Adjusted means
	(Yi)	(Xi)	(bw)	Yi—(bw) $(\bar{X}i-\bar{X})$
E-25	14.00	47.54	(.20)	14.01
E-50	13.09	38.54	(.20)	14.90
E-75	14.00	42.27	(.20)	15.06
E-100	14.54	47.91	(.20)	14.47
Control	23.36	61.64	(.20)	20.45
Grand Mean	15.80	47.58		

group (E-100) and comparing the highest experimental group's mean score on List 3 (E-100) with the lowest experimental group's mean score on List 3 (E-50) failed to reveal any significant differences among the experimental groups, even when a biased selection of means was made to allow differences to appear (E-25 vs. E-100: $t = 0.43$, *df*. 22, $p > 0.70$; E-100 vs. E-50: $t = 1.30$, *df*. 22, $p > 0.20$). Thus, although the combined experimental groups do differ from the control group they do not appear to differ significantly among themselves. In the absence of such differences, the significant over-all F was not convincing evidence for the presence of a mediation-effect.

The accompanying Fig. 1 presents a graph of the adjusted mean scores as an aid in visualizing the function which relates the scores obtained to the probability of the occurrence of the contingent mediator. The function is essentially linear except for the part of the curve between the control group and the E-25 experimental group. The flatness of the function bears out the findings of the *t*-test, previously reported, *i.e.* that few, if any, differences exist among the experimental groups which could be attributed to variations in the occurrence of the contingent mediator.

Following these basic analyses, several subsidiary analyses were conducted. The first analysis was conducted to test the hypothesis of a common regression line for

all groups. This hypothesis was found to be untenable (List 1–3, $F = 4.17$, *df.* 4, 49, $p < 0.01$; List 2–3, $F = 6.68$, *df.* 4, 49, $p < 0.01$; List 1, 2–3, $F = 3.49$, *df.* 4, 49, $p < 0.05$). From these significant *F*-ratios it can be concluded that the differences among groups were not due to differences in initial performance-level and that an effect due to the experimental treatment was present.

The second analysis was conducted to test the hypothesis of a common slope for the regression lines. This hypothesis was found acceptable (List 1–2, $F = .786$, *df.* 4, 45, *NS;* List 2–3, $F = 1.95$, *df.* 4, 45, *NS*; List 1, 2–3, $F = .09$, *df.* 4, 45, *NS*). These non-significant *F*-ratios indicated that although the regression lines were not common for all groups, as shown in the first subsidiary analysis, the slope of the regression lines for all groups was the same.

A third analysis, to test for linearity of regression, resulted in non-significant *F*-ratios for Lists 1–3 and combined Lists, 1, 2–3 (Lists 1–3, $F = 1.63$, *df.* 3, 49, NS;

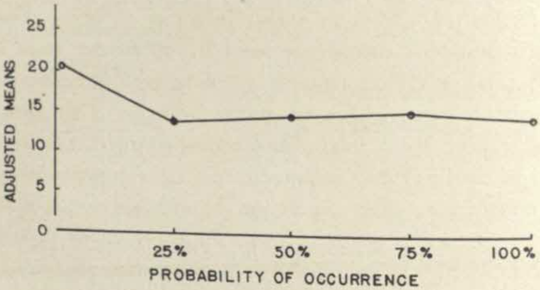


FIG. 1. ADJUSTED MEAN TRIALS TO CRITERION WITH LIST 3
(Controlled for initial performance on Lists 1 and 2.)

Lists 1, 2–3, $F = 1.80$, *df.* 3, 49, NS) but in a significant *F*-ratio for Lists 2–3 (Lists 2–3, $F = 7.68$, *df.* 3, 49, p. < 0.01). From these results it may be inferred that the regression is linear for Lists 1–2 and Lists 1, 2–3 but that the regression is probably non-linear for Lists 2–3. The non-linearity of List 2–3 may be attributed to the fact that in both the experimental and the control group some *S*s required more trials to reach the criterion of mastery for List 3 than for List 2. This would seem to indicate the presence of some negative transfer effects on List 3 as a result of learning new responses to old stimulus-objects.

Discussion

The statistical variance of our studies shows that a sensitive method of testing the effect of mediation was employed. Since we inferred no evidence of an effect due to mediation, we conclude that mediation either has no effect or its effect is so closely interwoven with general transfer that it cannot be distinguished even when fairly sensitive techniques are employed. If this conclusion is correct, then the postulation of separable mediation effects in verbal learning is untenable. Positive findings in

earlier studies of mediation can more satisfactorily be explained in terms of transfer. This explanation is more parsimonious and it is also more satisfactory, since it can account for observed results without recourse to theoretical constructs that lack experimental verification.

Effects of mediation have, to be sure, been demonstrated in a variety of contexts. A fact of considerable importance in considering this, however, is that these demonstrations have typically been within a transfer paradigm, either implicitly or explicitly. Furthermore, the broader context of the effects of generalization encompasses the concept of mediation. Mediation can be considered as a special case of transfer, and, to some extent, transfer can be considered as a special case of a more general theory of stimulus- and response-generalization. From this point of view, an explanation of certain facilitative effects in verbal learning as being wholly due to mediation has doubtful validity. It seems more to the point to consider that these facilitative effects (usually an increment in ease of mastery on criterional trials) involve general transfer-effects. This should not be construed as arguing for a total abandonment of mediation since it seems to be useful in dealing with response-mediated behavior involving physical response-sequences. The 'mediation hypothesis' (specifically that as formulated by Osgood) seems adequate in the context in which it was originally formulated.[8] Its extrapolation to cover the phenomena of transfer in verbal learning, where transfer and generalization effects may proceed along different dimensions than those involving physical response-sequences, seems, however, to be forced.

The establishment of a simpler conceptual framework for the interpretation of observed facilitative effects can best be approached by way of the existing concepts of transfer. Transfer is ubiquitous in learning studies and, as a consequence, it seems reasonable to consider some implications of transfer which will aid in understanding so-called effects of mediation. It is thought that more attention should be focused upon general transfer, broadly defined, in contradistinction to current emphasis upon limited specific stimulus-response relations.

In this connection, Harlow's work on the development of learning sets is *apropos*.[9] He indicates that the experimental emphasis upon comparatively specific and isolated responses, typified in many learning studies, has drawn attention away from more general and common situations in ordinary learning where behavior occurs in a broad context of stimuli and is repetitive in nature rather than being confined to isolated specific responses. From this general framework, he develops a theory that cumulative effects of repetitive experience within the learning situa-

[8] Osgood, *op. cit.*, 392-412.
[9] H. F. Harlow, The formation of learning sets, *Psychol. Rev.*, 56, 1949, 51-65.

tion results in a generalized set to learn. This situation is quite analogous to that which obtains with cumulative positive transfer in paired-associate learning. The generalized set to learn proposed by him would seem to be a result of some form of non-specific general transfer. Many of the observed facilitative effects found in studies of mediation in verbal learning are amenable to an interpretation in terms of this generalized effect of transfer.

One line of evidence for this position is offered in the theoretical analysis of transfer by Osgood.[10] In his treatment of transfer, a formal analysis of certain phenomena is set forth. Among other observations, he notes that the negatively accelerated curve of cumulative positive transfer could occur as a function of the learning of general habits from trial to trial. These habits, conceived to include the assimilation of general principles, modes of attack, or specific techniques, when transferred from trial to trial lead to a progressive facilitation of learning. Such a facilitation could be the basis for a postulated effect due to mediation although the effect may, in reality, be due to transfer. The advantage in ease of learning for experimental groups on criterional trials does not argue against this interpretation nor does this advantage indicate that the increment in facilitation of learning over and above that of the control group is due to mediation *per se*. It may more probably be due to an added enhancement of a generalized effect of transfer which is provided by alterations of the stimulus-conditions existing among the experimental groups. This interpretation of the experimental effect observed in studies of mediation is supported by the recent findings of Crawford and Vanderplas.[11]

With reference to our point that mediation does not exist independently of transfer, it should be noted that the design of the present experiment provides for the independent variation of only a single variable—the presence or absence of the mediating stimulus, all other factors being either controlled or their effect so randomized that no differential effects were present. The design, moreover, allowed for both positive and negative transfer. Since negative transfer is relatively transient, the facilitative effect of positive transfer remained dominant.[12] Equal amounts of negative transfer should have existed for both experimental and control groups, because of the use of identical learning tasks, although an alteration in the stimulating conditions for the experimental group could possibly have lessened the amount of negative transfer for this group.

A further examination of the conditions of this study indicate that some variations in the stimulating conditions were present, which produced the observed difference between the combined experimental groups and the control group on criterial trials.[13] Thus, changes in stimulating con-

[10] Osgood, *op. cit.*, 520-548.

[11] J. L. Crawford and J. M. Vanderplas, An experiment on the mediation of transfer in paired associate learning, *J. exp. Psychol.*, 47, 1959, 87-98.

[12] J. A. McGeoch and A. L. Irion, *The Psychology of Human Learning*, 2nd ed., 1952, 338.

[13] McGeoch and Irion, *op. cit.*, 425, 450; R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, rev. ed., 1954, 766.

ditions of the experimental groups (the contiguous presentation of the mediating stimulus-objects) may have decreased the amount of retroactive inhibition. If so, positive transfer would be enhanced by the reduction in negative transfer and would produce the apparent significant difference between the combined experimental groups and the control group.

SUMMARY

An evaluation of the relative contribution of mediation to transfer effects in verbal learning was undertaken in this study. The provision for observable mediating stimuli interposed in the presumed mediation process offered a sensitive comparison of mediated transfer as contrasted with general transfer. Statistical analysis suggested that there was little convincing evidence for the presence of an effect that could properly be attributed to mediation *per se*. It was thought that interpretation of the phenomena of mediation by recourse to the inferred action of mediating stimuli might more parsimoniously be ascribed to the effects of transfer and that, when general and specific transfer effects are controlled, the residual increment in learning attributable to mediation is insignificant.

MEASUREMENT OF ABSOLUTE OLFACTORY SENSITIVITY IN RATS

By WILLIAM R. GOFF, Yale University School of Medicine

Earlier researches on olfaction in animals have been qualitatively oriented, reflecting a lack of adequate methods for precise, quanitative measurement of sensitivity. This study, an attempt to supply this lack, describes a method of determining the absolute sensitivity of rats to three homologous hydrocarbons.[1]

The olfactometer described here controls the olfactory environment during stimulation. It permits the presentation of specifiable, homogeneous stimuli which are smelled under normal conditions of breathing and sniffing. Environmental control prior to and between stimulations is necessary as it eliminates general or selective adaptation which would result from exposure to irrelevant odors. The olfactometer presents the stimuli in the low concentrations that are necessary to measure absolute sensitivity.

The behavioral method of measuring sensitivity consists (1) in developing an operant discrimination in which presence of an odor in the response-chamber produces cessation of the operant response (bar pressing). When this has been established, the discrimination is then (2) made progressively more difficult by lowering gradually the concentration of the olfactory stimulus. The degree of discrimination is indicated by the number of bar-presses made in its presence.

METHOD

Apparatus. One- and two-liter Erlenmeyer flasks, fitted with Teflon stoppers,[2] serve as odor-containers (Fig. 1). Three types of flasks—generator, dilution, and reservoir—are used. The stopper of the generator-flask is pierced by a #18 hypodermic needle (*A*) with a valve (*B*) modified from a Luer Loc stopcock. A glass column (*C*), containing deodorizing chemicals, permits equalization of flask-pressure with the atmosphere by opening the stopcock (*D*). A mercury manometer (*E*) indicates flask-pressure. The dilution-flasks have only the hypodermic-needle-

* Received for publication March 23, 1960. This paper is adapted from a doctoral dissertation submitted to the University of Virginia. The author is indebted to Professor Frank A. Geldard, who directed the research.

[1] While this paper was in press, J. T. Eayrs and D. G. Moulton (*Quart. J. exp. Psychol.*, 12, 1960, 90-98) published an account of an apparatus for the quantitative measurement of olfactory acuity in rats.

[2] E. I. du Pont de Nemours and Company, Inc., Polychemicals Department, Wilmington, Delaware.

and-valve assembly and a glass stirrer (*F*) attached to a bar-magnet sealed in glass (*G*). The reservoir-flasks have a valve, manometer, stirrer, and glass-arms to connect the flask into the olfactometer via ground-glass joints. The arms may be sealed off with glass plugs (*I*).

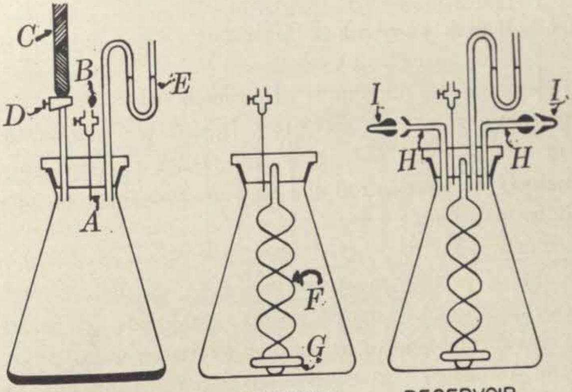


FIG. 1. THREE TYPES OF FLASKS

Odor-vapor is initially generated by allowing an odorant to vaporize in a generator-flask at constant temperature and atmospheric pressure until a vapor-pressure-to-partial-pressure equilibrium is reached. Under equilibrium-conditions, the concentration of the odorant can be calculated. The flask contains an excess of odorant. When odor-vapor is withdrawn, further vaporization of the odorant reëstablishes the equilibrium.

Odor-concentrations are lowered by successive dilutions. A dilution consists of withdrawing a sample from the generator and injecting it into a dilution- or reservoir-flask containing deodorized air. Since the maximal single-step dilution that is practicable is 2000 to 1 (1 cc. of odor-vapor injected into a 2-l. flask), odor-concentrations below approximately 10^{-2} mg./l., require intermediate steps using one or more dilution-flasks. The final reduction, however, is always made into a reservoir-flask.

Odor is transferred between flasks by means of Luer Loc hypodermic syringes. A syringe is locked onto the flask-valve, the valve opened, a measured amount of odor-vapor withdrawn, and the valve closed. A holding clip is attached to the syringe-plunger; the syringe is removed and locked onto the valve of the next flask. This valve is then opened, the holding clip removed, and odor injected. Since, prior to injection, an amount of air is withdrawn from the dilution-flask which equals the amount of odor to be injected, the contents of the flask are at atmospheric pressure after injection. Complete homogenization of the odor-vapor with the diluent air is critical. It is accomplished without violating the flask-seal by several minutes of rapid rotation of the glass stirrer by external magnetic activation.

When the necessary dilutions have been made, the reservoir-flask containing the

desired stimulus-concentration is connected into the main apparatus shown in Fig. 2. To present the stimulus to the animal, a two-way Luer Loc stopcock (A) is turned to connect the odor-reservoir to a 100-cc. hypodermic syringe (B). The syringe-plunger is raised to draw in 20 cc. of odor and is held momentarily in this position. The resulting vacuum in the odor-reservoir siphons purified water from a flask (C) until the entire system returns to atmospheric pressure, whereupon the siphon stops. The stopcock is then turned to connect the stimulus-syringe

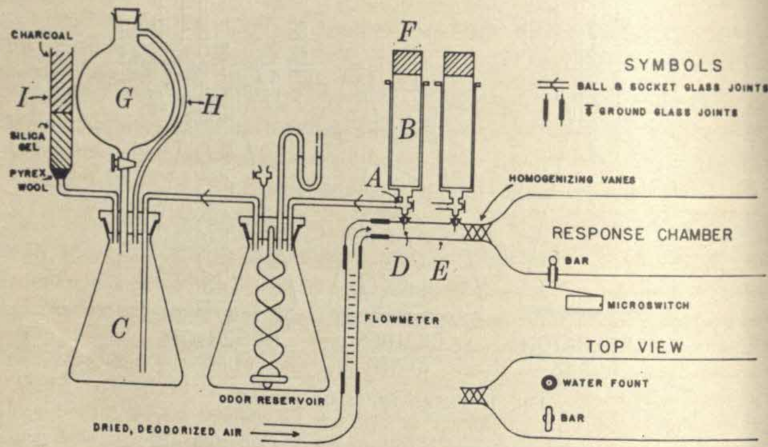


FIG. 2. SCHEMATIC DIAGRAM OF THE APPARATUS

to a hypodermic needle (D) which projects through a Teflon stopper into the tube (E) leading to the response-chamber. The plunger of the syringe is released and the weight (F) causes it to fall, injecting the odor into the air-stream flowing through the tube (E). The rate of odor-injection is 2 cc./sec. and the rate of air flow is 543.33 cc./sec. This further dilution is considered when calculating the stimulus-concentration. Three stationary vanes whose blades are set in opposition to each other homogenize the injected odor and flowing air. A second stimulus-system identical with that just described, permits the delivery of a non-odorous, control stimulus. The function of a stimulus-system, *i.e.* whether it presents odor or serves as the control, is varied randomly from day to day.

Repeated withdrawals from the reservoir without some means of pressure-compensation would progressively decrease the pressure and therefore the odor-concentration per unit volume. In the present experiment, the extremely low solubility of saturated hydrocarbons in water permitted replacing the stimulus-volume withdrawn with purified water. If water-soluble odorants are used, some other liquid in which the odorant is insoluble must be substituted. For equalization of pressure to be precise, the water level in (C) must be constant. Water flowing from (C) into the reservoir lowers the level in (C) admitting air into (G) via (H). (G) is a water-reservoir from which water flows to (C) until (H) is again sealed. All air admitted to this system passes through deodorizing chemicals in (I).

The olfactometer is easily disassembled. Prior to use, it is taken apart and cleaned in glass-cleaner followed by a hot sulphuricacid-potassiumdichromate solution. The dilution- and reservoir-flasks are flushed with deodorized air and sealed. Excepted from this daily procedure, the generator-flasks and their contents are kept continuously at equilibrium in a constant-temperature bath.

The response-chamber is a cylindrical glass 'wind-tunnel' 4 in. in diameter and 9½ in. long. A stream of air, dried and deodorized by passage through calcium chloride, coconut charcoal, and silica gel, flows through the chamber continuously. A glass response-bar and a drinking fount project into the chamber from the bottom. Water presented via the fount reinforces bar pressing. Responses are counted electrically. Cumulative records may be used to provide a graphic description of over-all behavior.

As an empirical check on the accuracy of the method of generating, diluting, and presenting the stimulus, odor-concentrations were independently determined using a gas chromatograph. The average error between the calculated odor-concentration and the concentration predicted on the basis of the chromatogram for concentrations of 0.208 mg./l. and 0.052 mg./l. was 5.56%.

Procedure. The procedure for developing the discrimination consists of four parts: (1) training the cessation of bar pressing to the odor-stimulus; (2) training continued pressing to the non-odorous stimulus; (3) testing the response to the odor-stimulus; and (4) testing the response to the non-odorous stimulus. The non-odorous, control-stimulus prevents any irrelevant cues which may be concomitant with the presentation of the odor-stimulus (*e.g.* sound of syringe-plunger falling, valves being turned) from becoming part of the discriminative situation and provides a baseline against which to compare responses to the odor.[8]

At the beginning of training, the bar-pressing response is shaped on a continuous-reinforcement schedule. As the response develops, the schedule is gradually shifted to a variable ratio of 3:1, 5:1, and, ultimately, 10:1. The animal's orientation in the apparatus while bar-pressing is critical. It is trained to press with its left forepaw while its nose is pointed into the oncoming air-stream. This position optimally orients the animal to receive the stimulus; odors from its body, from waste, as well as residual stimulus-odors, are flushed out of the chamber behind it.

When a reasonably stable rate has been developed at either 5:1 or 10:1, discriminative training is begun with a stimulus-concentration considerably above threshold. The odor-stimulus is presented for 10 sec. Reinforcement ceases with the onset of odor and remains off for a minimum of 45 sec. If, at the end of 45 sec., *S* has not touched the bar during the preceding 5 sec., reinforcement resumes. If this requirement is not met, reinforcement is withheld until 5 sec. of this type of withdrawal is obtained. The next press then is reinforced, and the regular schedule is resumed. The behavioral criterion for the resumption of reinforcement precludes reinforcing a period of pressing. This discriminative procedure is repeated at 1-2 min. intervals until presentation of the odor begins to produce

[8] A further check against the presence of irrelevant, discriminative cues may be introduced at the discretion of *E*. The odor-reservoir flask is removed from the system and replaced with a flask containing no odor. Training and testing are carried out with this pseudo-odor system using the same reinforcement-procedure as is used when odor is present. A cessation of bar-pressing when a pseudo-odor is presented indicates the presence of irrelevant cues.

cessation of response. At this point, the non-odorous control-stimulus is interposed, initially about every fifth stimulus, but with increasing frequency until odorous and non-odorous stimuli are being presented equally often. The duration of the injection of the control-stimulus during training is made sufficiently long to insure that the animal receives one or two reinforcements in its presence.

When the discrimination appears established, it is tested by comparing the number of bar-presses made in the first 10 sec. of odor-stimulation with the number made in 10 sec. of the control-stimulus with reinforcement absent. Since the latter measures the tendency to respond with all components of the odor-stimulus present except odor, a difference in responding from this baseline in the presence of odor is attributable to the odor.

Because repeated testing without any retraining would tend to destroy the discrimination, retraining trials are interspersed with testing trials throughout the course of sensitivity-measurements. The procedure for testing the reaction to the odor-stimulus is identical with the first 10 sec. of an odor-training trial; therefore, if the number of bar-presses made during the first 10 sec. of a training trial are noted, the training trial also serves as a testing trial.

In summary, training and testing the discrimination consists of the following: test-train odor, train non-odor control, and test non-odor control. A daily session includes each of these three occurring five times in random order. The discrimination is considered established when 5 or fewer responses occur in the presence of odor in contrast to 20 to 30 responses in the presence of the non-odor control.

When the discrimination is established, the sensitivity to a given odor is measured by progressively decreasing the stimulus concentration by log steps while continuing the training and testing procedure. The discrimination is trained and tested at each concentration for three days. The absolute threshold is defined as the point of discriminative failure. Discrimination is judged to have failed at that concentration where the same number of responses are emitted to odor as to non-odor for three successive days. In other words, the absolute threshold is the strongest odor-concentration at which no difference in the animal's behavior toward the odorous and control-stimuli can be measured.[4]

RESULTS

In previous work, the odor-stimulus was present during the entire odor-training period.[5] In the present work, the discrimination was established just as readily with odor present only during the first 10 sec. of this period. Using a brief 'puff' instead of a prolonged stimulus reduces the possibility of adaptation to repeated stimuli and reduces post-stimulus evacuation-time. Furthemore, smaller volumes of odor per animal are required.

---

[4] The threshold-criterion of Eayrs and Moulton (*op. cit.*, 90-98) is comparable to that used here. In their two-choice situation, threshold is taken as the concentration yielding 50% correct response, or, in other words, where there is no difference in response to the two stimuli.

[5] Carl Pfaffmann, W. R. Goff, and J. K. Bare, An olfactometer for the rat, *Science*, 128, 1958, 1007-1008.

The absolute sensitivity of several male albino rats to the odors of *n*-Pentane (C_5H_{12}), *n*-Hexane (C_6H_{14}), and *n*-Heptane (C_7H_{16}) was measured. Fig. 3 presents the results for pentane. The ratio of the number of responses made in the presence of the odor-stimulus to the number made in the presence of the control-stimulus is plotted on the ordinate.

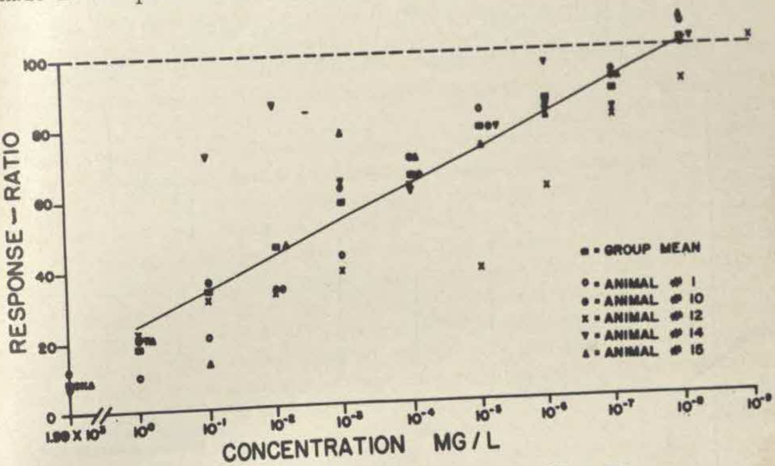


FIG. 3. DISCRIMINATIVE FUNCTION FOR *n*-PENTANE

The abscissa represents log stimulus-concentration in milligrams of odorous substance per liter of total air-odor mixture passing through the response-chamber. Each point on the graph is the mean of three days of measurement or a total of 15 comparisons.

Alexander's "Trend Test" was applied to the date.[6] The results showed that changes in the response-ratio resulting from changes in concentration were significant beyond the 0.1% level and that the change is a linearly increasing function of log-concentration decrement. The best fitting straight line in Fig. 3 accounts for 98.16% of the variance of the function $y = -15.76 + 9.46\,x$. There were no significant differences between animals, and no significant interactions between animals and concentrations.

To test the adequacy of a three-day period of response-measurement at each concentration, the effect upon variability of prolonged measurement at the same concentration was determined. Two animals (#14 and #15) who showed the upper and lower extremes of variability were tested for 15 additional days at a concentration of 10^{-5} mg./l. pentane. If these variabilities reflected fluctuations in olfactory sensitivity, they should remain es-

[6] H. W. Alexander, A general test for trend, *Psychol. Bull.*, 43, 1946, 533-557.

sentially unchanged with continued measurement at the same concentration. If the difference in the two animals' variabilities reflected a difference in degree of discriminative learning, the animal with high variability should become more stable with additional training. Finally, if variability is due primarily to chance-factors, it should regress toward an intermediate value for both animals. The results fulfilled this last condition. The variability of the response-measure thus seems due to chance-fluctuations, and group-

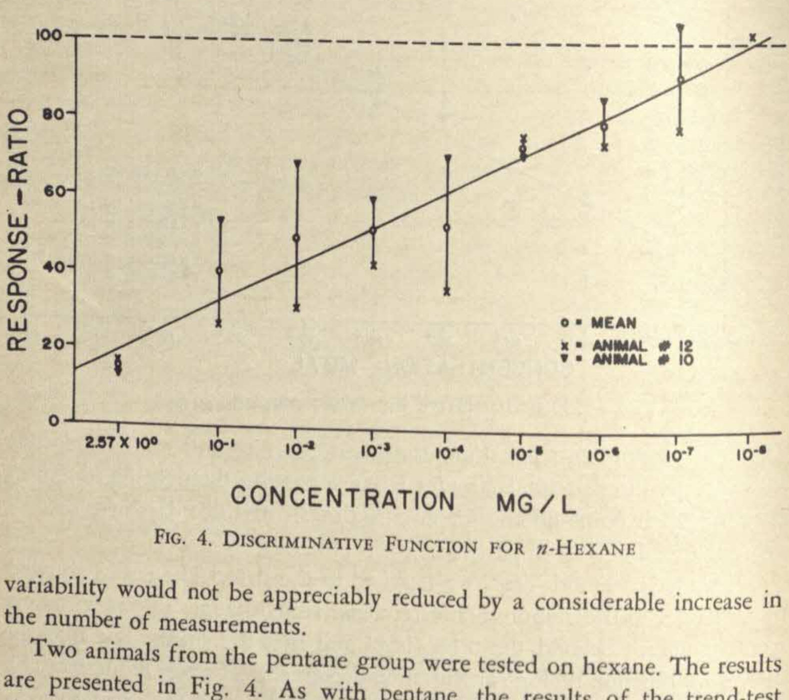


FIG. 4. DISCRIMINATIVE FUNCTION FOR *n*-HEXANE

variability would not be appreciably reduced by a considerable increase in the number of measurements.

Two animals from the pentane group were tested on hexane. The results are presented in Fig. 4. As with pentane, the results of the trend-test showed differences in response-ratio as a function of changing concentration to be significant beyond the 0.1% level and the data adequately described by a straight line. The best fitting straight line accounts for 93.75% of the variance of the function $y = 13.34 + 9.64\,x$. There were no significant differences in effect of concentration upon the two animals. Between individual animals, however, there was a difference significant at the 1% level. Fig. 4 shows that animal #12 generally responded less during odor-stimulation. Furthermore, concentration had to be lowered an additional log step to produce discriminative failure in this animal. Reference to the pentane data (Fig. 3) shows that this same animal gen-

erally responded less, and that its discrimination also failed one log unit lower with pentane than the other *S*s.

Fig. 5 presents the results of measurements of sensitivity to the odor of heptane for 2 *S*s. These animals were among the first run in this experiment, and the data were not collected in the systematic form later used. The numbers in brackets give the number of days that the particular concentration was tested. Numbers in parentheses give the order in which the

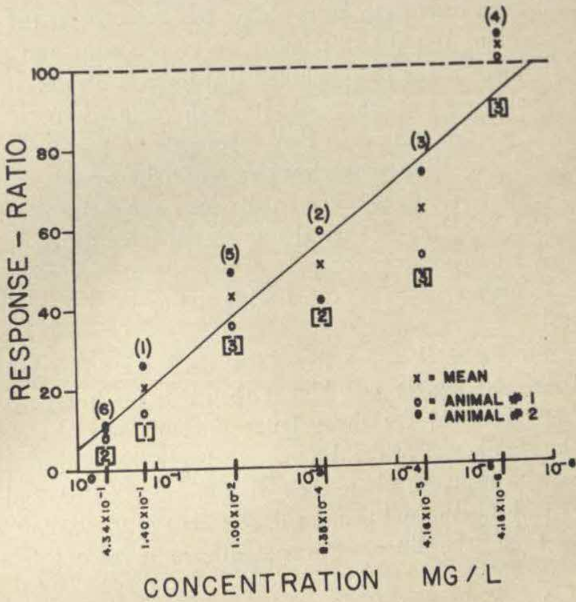


FIG. 5. DISCRIMINATIVE FUNCTION FOR *n*-HEPTANE

concentrations were tested. Because of the unequal concentrations, steps, and numbers of measurements, the trend-test was inapplicable. The straight line is fitted by the method of least squares. These results are particularly interesting because they indicate a regular increase in the rate of response to the odor-stimulus as concentration is decreased regardless of the sequence in which concentrations are tested.

The absolute threshold has been defined as the first odor-concentration in a progressively decreasing series which the animal does not discriminate from the control-stimulus. In a discrimination involving a periodically reconditioned operant response, discrimination is observed as any stimulus-

evoked deviation from a previously established constant rate of response.[7] A schedule of variable ratio reinforcement produces such a constant rate after adequate training. Furthermore, in the present experiment, a basal rate of response is actually measured under control conditions which duplicate odor conditions except for the presence of odor. The response-ratio is the comparison of odor-response-rate to control-stimulus-baseline-rate expressed as a percentage. Thus, a response-ratio significantly different from 100 indicates that the odor is discriminable.

Inter-odor comparisons can be made on this basis. If, at the same physical concentration, *S*s show discrimination of one odor and fail to discriminate another, it may be concluded that there is greater sensitivity to the discriminable odor. With regard to such comparisons, it should be stated that it was impractical in this experiment to test large groups of animals. Valid inter-odor comparisons of sensitivity would require, therefore, that the same animals be tested on all odors, with each animal serving as its own control. The death of several animals prevented the complete fulfillment of this requirement in the present experiment. Comparisons of the remaining animals must be regarded with caution because of this incomplete control.

Reference to Fig. 3, 4, and 5 shows that the average threshold for heptane ($4{:}16 \times 10^{-6}$ mg./l) was higher than that for hexane or pentane. The average threshold for these latter odorants was the same (10^{-8} mg./l). The significance of the difference between the mean response-ratio of the 2 *S*s with heptaine, where their discrimination failed, and the mean of the 7 *S*s with hexane and pentane *at this same concentration* was determined by a *t*-test. The difference was significant at the 2% level of confidence.

Fig. 3 and 4 show no difference between the mean thresholds for pentane and hexane, but 2 *S*s were common to both substances and could be directly compared. Furthermore, as has already been stated, the data indicate that one of these animals was more sensitive than the other to the odors of both substances. It seemed possible that for these two animals, at least, grouping the data obscured between-odors differences. To make a direct comparison, the significance of the difference between the mean response-ratios of each animal at the concentration where its discrimination of hexane failed and the mean response-ratio for pentane at the same concentration, corrected for the difference between animals, was treated

[7] B. F. Skinner, *The Behavior of Organisms,* 1938; F. C. Frick, An analysis of an operant discrimination, *J. Psychol.,* 26, 1948, 93-123.

by a t-test. The difference was significant beyond the 1% level of confidence, suggesting that, at the concentration where the animals failed to discriminate the odor of hexane, they were significantly more sensitive to pentane.

Summary

An olfactometer has been developed which measures the absolute olfactory sensitivity of rats. The thresholds of a small number of animals for the odors of *n*-Pentane, *n*-Hexane, and *n*-Heptane were measured. In this operant discriminative situation, the threshold is defined as the strongest concentration at which no difference in the animal's behavior toward odorous and control-stimuli can be measured. As odor-concentration is decreased from suprathreshold training levels to the point of discriminative failure, the rate of response to the odorous stimulus compared to a control-stimulus rate was found to be an increasing linear function of logarithmically decreasing concentration. The data of this experiment suggest that, for these chemically homologous odorants, the threshold for heptane was highest and that for pentane lowest.

PRINCIPLES OF MOMENTUM AND KINETIC ENERGY IN THE PERCEPTION OF CAUSALITY

By THOMAS NATSOULAS, Wesleyan University

By means of a series of laboratory experiments, Michotte has attempted to specify some of the stimulus-conditions which produce various perceptions of causality.[1] Using small, painted squares which appear to move horizontally and to make head-on contact, Michotte demonstrated that the kind of reports of causation which *S*s give in response to the 'collision' of the two squares depends on, among other variables, the relative velocities of the two squares.[2]

The present report proposes two alternative principles for purposes of ordering reports of causation and describes an experimental test of the two principles. The type of stimulus-situation studied is one of the simplest Michotte has used: a black square (A) moves in a straight horizontal line at a uniform velocity toward a red square (B), makes head-on contact with it, and becomes stationary while B moves away from A in the same straight line. Both of the proposed principles predict from the apparent velocities and masses of A and B (V_A, V_B, M_A, and M_B) to the reports of causation which are given to the particular stimulus-situation.

According to the principle of psychological momentum, when $M_A V_A / M_B V_B = 1$, *S*s will report that A alone produced B's movement, that A propelled B, gave B a push, that A's collision with B is the only source of force for B's movement (launching). If this ratio of momenta is less than one, *S* will report that A released B's movement, triggered it off, made it possible but did not produce it (releasing). Finally, if the ratio exceeds unity, A will be reported as having launched B, but that B's movement seemed to be impeded by an additional force (braking).

The principle of psychological kinetic energy makes similar predictions: if $M_A V^2_A / M_B V^2_B \geq 1$, responses of launching are expected; if this ratio is less than one, then reports of releasing should occur. This principle makes no prediction to reports of braking. The rationale for the absence of such a prediction derives from

* Received for publication March 28, 1960. This article is based on a dissertation submitted to the Department of Psychology of the University of Michigan in partial fulfillment of the requirements for the Ph.D. degree. The author is indebted to Dr. David Birch for his assistance.

[1] Alberto Michotte, *La perception de la causalité*, 2nd ed., 1954, 1-306, esp. 41-143.

[2] Michotte, *op. cit.*, 104-109.

physics. A loss of kinetic energy in any actual collision does not imply hindrance of *B*'s movement by another force. All physical collisions of bodies of known substances involve a loss of kinetic energy in deformation and heat. It is further assumed that *Ss*' reports will always be in terms of the fewest sources of force possible. According to the priniple of kinetic energy, no responses of braking will, therefore, occur. Launching, which involves only one source of force, will suffice when there is a decline in kinetic energy subsequent to the collision. Michotte makes no reference to a response of braking in his studies.[3]

Strictly speaking, the principles introduced here predict from *S*'s own judgments of velocities and masses to his reports on perceived causality. The predictions derived from the principle of psychological momentum are: (a) the greater the mean value of $M_A V_A/M_B V_B$ for a stimulus-condition, the smaller the number of *Ss* who will respond with reports of releasing to that condition, and the greater the number of *Ss* who will report braking, and (b) the more the mean value of $M_A V_A/M_B V_B$ departs from unity in either direction, the smaller the number of *Ss* who will report launching. From the principle of kinetic energy, it follows that the greater the mean value of $M_A V_A^2/M_B V_B^2$ for a condition, the smaller the number of *Ss* who will report releasing, and the larger the number who will respond with reports of launching.

METHOD

Stimulus-conditions. All the stimulus-conditions used had the following properties in common: *A*, a black square or rectangle, appears near the left end of a horizontal aperture, 0.5 cm. in height (vertical) and 11.0 cm. in width (longitudinal). Simultaneously *B*, a red square or rectangle, appears at the center of the aperture, 4.5 cm. from *A*. For 0.27 sec. both *A* and *B* are stationary at their original positions. *A* then commences to move toward *B*. When *A* comes into contact with *B*, they remain in stationary contact for 0.03 sec. Then *B* crosses 4.5 cm. to the right, at which point the aperture ends and *B* disappears behind the screen. *A*, too, disappears at its stationary position. This sequence occurs 10 times for every stimulus-condition on every trial, the 10 presentations taking place in 15.0 sec.

The stimulus-conditions vary with respect to the velocity-ratios of *A* to *B*. With velocities of 60, 40, 20 cm./sec., five ratios are produced: 3:1, 2:1, 1:1, 1:2, 1:3. For example, a velocity-ratio of 3:1 means that *A* moves to meet *B* at 60 cm./sec. and that *B* moves off after contact at 20 cm./sec. *A* and *B* also vary with respect to size in the same five ratios. They have a width of 0.5, 1.0, or 1.5 cm. and a height of 0.5 cm. A size-ratio of 1:2 indicates that *A* is a 0.5-cm. square while *B* is a rectangle, 0.5-cm. high and 1.0-cm. wide.

Apparatus. A description of the apparatus and of the way in which the stimuli are constructed is given in another article.[4]

---

[3] Michotte, *op. cit.*, 1-306.

[4] Thomas Natsoulas, Judgments of velocity and weight in a causal situation, this JOURNAL, 73, 1960, 404-410.

Subjects. The *Ss* (19 men) were students in elementary psychology at the University of Michigan. They were required to participate in a psychological study and were selected at random from sections not using a text that contained a discussion and illustrations of Michotte's studies.

Procedure. The *Ss* participated individually (three cases) or in pairs. Those in pairs sat side by side, with a thin partition of cardboard preventing them from seeing each other or each other's written responses. All *Ss* sat 5 ft. from the apparatus with the aperture at eye-level. The following instructions were given.

Instructions. This is a study in the perception of movement. You will see in this aperture two small squares or rectangles, *A* (a black one) and *B* (a red one). They will move across the aperture at various velocities and will seem to possess various weights. *A* will move from this end of the aperture to the center, where it will meet *B*. *B* will then move from the center to the right and disappear behind the screen. They will then reappear at their original positions and then go again through the same movements as before. Your task for each of the 75 stimulus-conditions is to report how you experience each of them. This will be made easy for you as a result of a study we made of just such impressions. We have managed to boil them all down to the three categories on these sheets. [At this point each *S* received a sheet which read as follows]: In each case choose (I), (II), or (III). Choose the one which best fits your impression.

(I) *A*'s movement produces *B*'s movement directly by means of their impact. *B*'s motion is due entirely to *A*'s motion and its collision with *B* and nothing else. The only source of force for *B*'s movement is *A*'s movement and *A*'s impact on *B*.

(II) *A*'s movement produces *B*'s movement directly by means of their impact—just as (I)—*but* some other force seems to prevent *B* from moving away from the point of impact with as much force as you would expect in view of the collision with *A*.

(III) When *A* contacts *B* it releases *B*'s movement. *A* thus allows *B*'s movement to take place and does not directly produce *B*'s movement. *B*'s movement seems to be due to a force released by *A*'s contact with *B* and not to a force produced by collision.

[The oral instructions then continued]: When each stimulus is presented you are to decide which of these three descriptions best fit your impression of what is happening. On the sheet before you, enter either (I), (II). or (III) as the case may be. You are not required to use the three categories equally often, and the order in which they are listed has no significance. Keep the sheets in sight and consult them from time to time to make sure you have the right one paired with the right number. There will be 75 stimulus-conditions for you to respond to, and you will see each of them occur 10 times. Watch each carefully; you will have plenty of time to write down your responses while I change the conditions. At the end of each set of 25 we will have a 5-min. break. Now let us read over the three impressions together.

The three alternative responses were read aloud and questions were answered.

The 25 stimulus-conditions were presented three times, each time in a different order. All three orders were counterbalanced with respect to the two stimulus-dimensions. Within each successive group of five trials, every ratio of velocities and every ratio of sizes occurred only once. Furthermore, no two stimulus-conditions having either the same velocity- or size-ratio were given in succession.

RESULTS

Physical variables. Fig. 1a illustrates the relationship found between velocity-ratios and the frequencies of responses in each causal category.

Trials 2 and 3 are so combined that every *S* is represented 10 times at every value of the abscissa; every *S* responded twice to each of the five stimulus-conditions having the same velocity-ratio of *A* to *B* and differing in the size-ratios of *A* to *B*. The frequency with which braking was reported increased and the frequency with which reports of releasing occurred decreased as the velocity-ratios favored *B* less and *A* more. Launching was most often reported when the velocity-ratio was 1 : 1. Departures from this ratio in either direction resulted in a decline in the number of

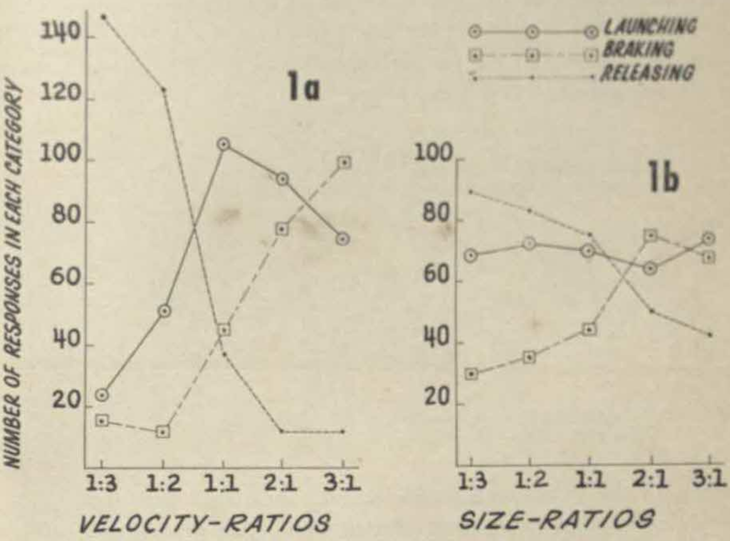


FIG. 1. RELATIONSHIP BETWEEN CAUSAL REPORTS OF VELOCITY- AND SIZE-RATIOS (Trials 2 and 3)

reports of launching. This decline was steeper when *A*'s velocity was less than *B*'s than it was when the reverse was true.

Fig. 1b illustrates the manner in which the size-ratios are related to the frequencies of response. The frequency with which releasing was reported declined and the frequency of braking, with one exception—from ratios of 2 : 1 to 3 : 1—increased as the size-ratios of *A* to *B* favored *B* less and *A* more. Launching does not appear to be related in any simple fashion to the size-ratios, the line in Fig. 1b being very nearly horizontal.

Table I summarizes a statistical evaluation of the above relationships, performed by applying to the data a multivariate analysis of information-

transmission.[5] In all analyses, Trial 1 has been treated as a practice-trial and is not included.

An examination of Table I reveals that there are three statistically significant sources of information: the velocity-ratios (without regard to variations in size-ratio), the size-ratios (without regard to variations in velocity-ratios), and the joint classification of the stimulus-conditions based on both velocity- and size-ratios.

Ratios of momentum and kinetic energy. In another phase of the present study, two groups of 19 *S*s each, chosen in the same way and from the same population as the *S*s of the first part of the study, responded under the same situations as the present experimental group to each of the 25 stimulus-conditions.[6] One group's task was to judge the conditions with

TABLE I

INFORMATION-TRANSMISSION ANALYSIS OF RESPONSE-FREQUENCIES

Source	*df.*	$-2 \text{ Log}_e \Lambda$	p*	% uncertainty accounted for
Velocity-ratios(V)	8	443.94	<.001	21.4
Size-ratios(S)	8	60.53	<.001	2.9
Trials(T)	2	3.21	.10–.20	0.2
V & S	32	93.25	<.001	4.5
V & T	8	12.88	.10–.20	0.6
S & T	8	2.05	.95–.98	0.1
V, S, & T	32.	41.27	.10–.20	2.0

* Chi-square test.

respect to the velocity-ratios of *A* to *B*. The second group judged the weight-ratios of *A* to *B*. The geometric means across *S*s judging velocity-ratios (V_A / V_B) and those judging weight-ratios (M_A / M_B) were used for purposes of calculating the ratios of psychological momentum ($M_A V_A / M_B V_B$) and kinetic energy ($M_A V_A^2 / M_B V_B^2$) for each of the 25 conditions.

Tau-coefficients have been calculated between the rankings of the conditions based on the two theoretical ratio-dimensions and the rankings based on the frequencies of response in each category of causation. Table II presents the values of *tau* obtained. Also included in Table II are values of *tau* based on rankings of the stimulus-conditions according to the geometric means of the judged weight-ratio and of the judged velocity-ratios. The *tau*-coefficient used is the one Kendall calls tau_a in his

[5] W. J. McGill, Multivariate information transmission, *Psychometrika*, 19, 1954, 97-116.

[6] Natsoulas, *op. cit.*, 405-407.

discussion of tied ranks.[7] It involves no correction for the presence of conditions tied in rank on either dimension. Tied ranks occur only for the dependent rankings. Because the predictions to reports of launching are such as to require that the number of launchings reported is greatest at the 1:1-point of a theoretical dimension and declines with departures in either direction from that ratio, for purposes of calculating *tau*-values, the hypothetical dimensions have been 'folded' at their 1:1-points in such a way that a momentum-ratio of 1:2.5, for example, is treated as equal to one of 2.5:1.

It can be seen in Table II that, with the exception of the geometric means of the judged weight-ratios, all variables are negatively correlated with frequencies of reports of releasing and launching and positively

TABLE II

TAU-COEFFICIENTS BETWEEN THEORETICAL RANKINGS OF THE STIMULI AND RANKINGS BASED ON FREQUENCIES OF RELEASING (R), LAUNCHING (L), AND BRAKING (B)

Theoretical rankings	Rankings based on response-frequencies					
	Trial 2			Trial 3		
	R	L	B	R	L	B
Kinetic energy	−.770	−.733	+.753	−.783	−.670	+.700
Momentum	−.617	−.413	+.540	−.597	−.417	+.607
Velocity	−.557	−.533	+.607	−.603	−.397	+.520
Weight	+.023	+.193	−.080	+.097	+.270	−.013

(*Tau*s of absolute magnitude greater than 0.365 are significantly greater than zero at the 1% level.)

correlated with the frequencies of reports of braking. Weight-ratios are correlated positively with frequencies of reported releasing and launching, and negatively with frequencies of reports of braking. The ratios of kinetic energy correlate more highly with the frequencies of reports of launching, braking, and releasing than do any of the other three predictive variables.

To evaluate the differences in effectiveness of the four predictive variables listed in Table II, an analysis of variance was performed on the absolute magnitudes of the *tau*-coefficients in Table II. The main effect of the four theoretical orderings of the stimuli is found to be statistically significant ($F = 159.77$) $df = 3$, $P < 0.001$) when tested against the interaction of theoretical orderings and trials ($df. = 3$). The only other significant F-value is found for the interaction of theoretical orderings and orderings of the stimuli based on frequencies of releasing, launching,

[7] M. G. Kendall, *Rank Correlation Methods*, 1948, 36-37.

and braking ($F = 6.56$, $df. = 6$, $P < 0.05$) when tested against the three-way interaction of theoretical orderings, orderings based on frequencies of response and trials ($df. = 6$). An F-test comparing the sums of the absolute *tau*-values obtained under the 'condition' of ordering the stimuli according to their ratios of kinetic energy and the sums obtained under an ordering based on ratios of momentum resulted in statistical significance ($F = 48.03$, $df. = 1$, $P < 0.01$) when tested against the interaction of theoretical orderings and trials.

Discussion

The psychological principle of kinetic energy, as defined here, predicts best the *S*s' reports of all three types of causation. Yet one type of report, braking, was not expected to occur at all according to this same principle. The basis of such a prediction was that a decline in kinetic energy from before to after the collision is a natural consequence of any collision and does not require the experience of an additional force hindering *B*'s progress. Michotte's results are more in keeping with the prediction of no reports of braking in that he does not indicate that such a report was given by his *S*s. A comparison of his procedure with that of the present experiment suggests a possible resolution.

Michotte's *S*s were permitted to describe their impressions spontaneously; the present procedure required *S*s to choose a description to fit their experience. The prediction of no reports of braking from the psychological principle of kinetic energy may be correct in the spontaneous case and not in the case where a choice between alternative responses is required. Thus, the *S*s will not utilize spontaneously categories involving a larger number of sources of force, but will do so when such responses are made available to them. Consequently, the category of braking is utilized in the present experiment despite the fact that the principle which seems to order the results best does not predict its utilization, and despite the fact that it involves an additional force in describing *B*'s apparent activity while the category of launching does not.

The validity of this interpretation is supported by an analysis of the data in terms of reproducibility, *i.e.* the extent to which the stimulus-conditions used in the present experiment constitute a unidimensional scale with respect to the reports of causation.[8] When reports of releasing were treated as the positive response, there was some, though not strong, evi-

---

[8] Natsoulas, A study in the perception of causality, unpublished Doctoral dissertation, University of Michigan, 1959, 44-45.

dence for the existence of a unidimensional scale of the conditions. On the other hand, treating the reports of braking as positive responses resulted in evidence for a scale that was negligible. The results indicate less consistency in the use of the category of braking as compared with that of releasing. Such inconsistency would arise when the *S*s do not check against the list of responses provided and react spontaneously, as in Michotte's studies, with a choice of launching.

With respect to the *tau*s of Table II, there is one particularly noticeable effect, that of the weight-ratios. The signs of the correlations which involve this variable are opposite to those found with the other predictive variables. Light may be shed on this disparity by returning to the psychophysical phase described earlier.[9] It will be recalled that two other groups of 19 *S*s responded to the same conditions used in obtaining reports of causation. One group's task was to judge the weight-ratio of *A* to *B* for each condition. It was found that the geometric means of the judged weight-ratios were directly related to the physical size-ratios and inversely related to the physical velocity-ratios. Thus, the judgments of weight may be thought of as comprised of two components: (a) a component, determined by physical size, which is inversely related to the frequencies of reports of launching and releasing and directly related to the frequencies of reports of braking, just as the remaining variables of Table II are; and (b) a component determined by the inverse relationship to velocity and therefore inversely related to whatever is directly related to the velocity-ratios. From a combination of these two factors may come correlations close to zero, and even in the direction opposite the correlations involving the other variables.

SUMMARY

An experimental test was conducted of two principles in the perception of collisions, a psychological principle of momentum and one of kinetic energy. Nineteen men (college students) were presented with 25 experimental collisions on each of three trials, each time in a different counterbalanced order. They were required to choose from descriptions of launching, braking, and releasing, that description which fitted each collision best. All collisions involved impact between two figures, one of which (*B*) was stationary prior to the collision and the other (*A*) was stationary after the collision.

On the basis of the judgments of two other groups of comparable *S*s,

[9] Natsoulas, *op. cit.*, this JOURNAL, 73, 1960, 407-409.

each of the stimulus-conditions was given a geometric mean value for judged velocity-ratios (V_A / V_B) and for judged weight ratios (M_A / M_B). From these values a ratio of momenta, ($M_A V_A / M_B V_B$), and a ratio of kinetic energies, ($M_A V_A^2 / M_B V_B^2$), were calculated for every condition.

According to the psychological principle of momentum, if the ratio of momenta is unity, launching of *B* by *A* should be reported. If the ratio is greater than one, braking should be reported, and if it is less than one, releasing should be reported. According to the psychological principle of kinetic energy, if the ratio of kinetic energies is equal to, or greater than, one, launching is the expected report, and if the ratio is less than one, releasing will be reported.

Orderings of the stimulus-conditions in accordance with either ratios of kinetic energy or ratios of momentum correlated significantly with orderings of the conditions in terms of frequencies of each of the three types of reports of causation. The *tau*-coefficients for the ordering of the conditions according to kinetic energy were greater in all cases than those found for the ordering according to momentum.

TASTE-DISCRIMINATION IN THE MONKEY

By J. S. SCHWARTZBAUM, University of Wisconsin, and
W. A. WILSON, JR., Bryn Mawr College

Taste-discrimination in animals has been investigated traditionally by the use of some variant of the 'preference method' in which the discrimination is evidenced by the relative amount of a fluid that is consumed in a free-choice situation. The limitations of this procedure as a way of determining taste-sensitivity are, by now, well organized.[1] It confounds the discriminative function of the taste-stimuli with a reinforcing function, since the reinforcement is obtained directly from the solutions to be discriminated. Consequently, it becomes uncertain whether a failure to discriminate relates to factors of taste-sensitivity or to the reinforcing role of the stimuli, or both. Moreover, under conditions which involve the consumption of appreciable amounts of hypertonic fluids, taste and preference become further confounded with post-ingestional osmotic factors as determinants of the discriminative behavior.[2]

A method has been employed to study taste-sensitivity in rats, which, to a certain extent, separates the discriminative from the reinforcing functions of the taste-stimuli.[3] Animals that had been deprived of water were forced to discriminate between two fluids which they consumed, in order to avoid a noxious electric shock. The significance of this change in the experimental conditions is seen in the finding that, although adrenalectomy lowers the preference-threshold for salt (NaCl), it does not increase the absolute taste-sensitivity to salt.[4]

In the present experiment a nonpreference method for studying taste-discrimination in monkeys was explored. Animals that had been deprived

---

* Received for publication March 27, 1959. This study was conducted at the Research Laboratories of the Institute of Living, Hartford, Connecticut, under a grant, No. M-546(C), from the U.S. Public Health Service, National Institute of Mental Health.

[1] Carl Pfaffman, Taste and smell, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 1143-1171.

[2] J. A. Carpenter, Species differences in taste preferences, *J. comp. physiol. Psychol.*, 49, 1956, 139-144.

[3] W. J. Carr, The effect of adrenalectomy upon the NaCl taste threshold in rat, *ibid.*, 45, 1952, 377-380; A. E. Harriman and R. B. MacLeod, Discriminative thresholds of salt for normal and adrenalectomized rats, this JOURNAL, 66, 1953, 465-471.

[4] Carr, *op. cit.*, 377-380; Harriman and MacLeod, *op. cit.*, 465-471.

of food were required to discriminate between two solutions differing in taste-properties to obtain a separate food-reward. The experiment was intended to provide information concerning: (a) the ability of monkeys to acquire such a discrimination; (b) their taste-threshold for NaCl; and (c) two related methodological problems.

METHOD

Subjects. The *Ss* were eight rhesus monkeys, five males and three females, of pre-adolescent age. Their body-weights ranged from 5.0 to 6.5 lb. All animals had prior training on paired-comparison tests of preferences for different kinds and amounts of foods. The *Ss* were maintained on a daily diet of 8–10 Purina Laboratory Chow pellets, supplemented by one-fourth of an orange. They were deprived of food for approximately 22 hr. before every test. Water was freely available in the home quarters.

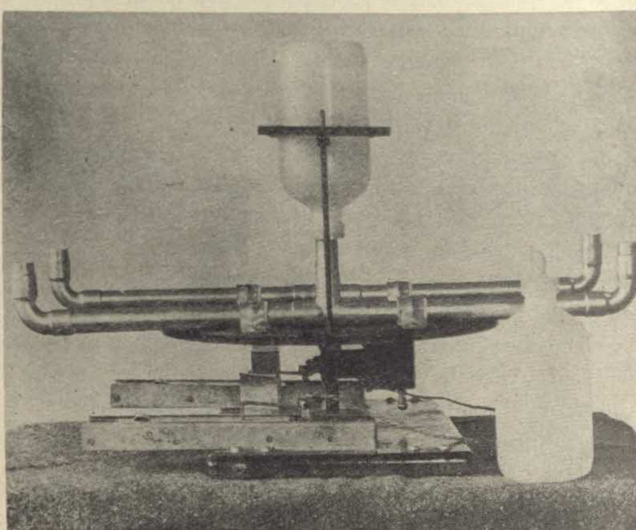


FIG. 1. MODIFIED WISCONSIN APPARATUS FOR TASTE DISCRIMINATION

Apparatus. All tests were conducted in an air-conditioned, sound-proofed room, using a modified form of the Wisconsin apparatus for taste-discrimination. In Fig. 1 a picture of the apparatus is shown. The taste-stimuli, distilled water and a solution of NaCl in distilled water, were contained separately in the two symmetrically designed tubes. These tubes, measuring 20.75 in. from outside end-to-end, were made of copper with inside diameters of 11/16 in. joined together by sweated fittings. The height of the right-angle, end sections was 2.25 in., and that of the T-shaped, center section was 2.75 in. The exposed surfaces of the tubes were sprayed with an aluminum metallic paint to provide a more even finish.

The fluid was supplied to each of the tubes from a 500-cc. polyethylene bottle

which was inverted over the T-shaped section through a supporting structure. A plastic cap which screwed on to the bottle was fitted with a piece of polyethylene tubing 0.5 in. in diameter. The plastic tubing extended 0.62 in. into the stem of the T-section, that the level of fluid in the right-angle, end sections of the copper tube would be maintained within 0.12 in. of the top. If the fluid fell below this level, air entered the plastic tube and displaced additional fluid from the bottle. A small hole was drilled near the top of the T-section for this purpose.

The metal tubes were mounted parallel to one another 6.75 in. apart on a turntable 11.75 in. diameter. The turntable, which was finished in a flat black, could be easily rotated to reverse the spatial positions of the tubes. Beneath the turntable, mounted on the vertical bar, was a spring-lock which fitted into either one of two holes that were drilled in the turntable. Thus, the turntable could be locked in one of two alternate positions for presenting the stimuli. Withdrawing the lock freed the turntable that it could be rotated.

Beneath each tube facing *S* a food cup with a moveable lid was enclosed which *S* was taught to retract. In Fig. 1 the lids are seen in the open position. A metal L-shaped rod was so mounted on a pivot behind each lid that the lid could be locked in the closed position. *E* controlled the unlocking of the lids by means of a cord attached to the rear of each rod. This action was accompanied by a distinctive clicking sound from the impact of the rod on an overhanging metal plate. The heavy metal base of the apparatus was so fitted with runners that the whole unit could be moved smoothly along a slide.

In addition to a moveable opaque screen in the Wisconsin box, which shielded the apparatus from *S* between trials, a one-way screen made from gelatin film and 0.25-in. Lucite panels was mounted on *S*'s side of the opaque screen. The one-way screen contained slots through which the stimuli to be discriminated and food cups were presented to *S*. The tubes projected 2.25 in. into the test cage at a distance of 10.12 in. above the floor of the cage. The food cups, which were spaced 6.87 in. apart, also projected between the vertical bars in the front of the cage. Two appropriately shielded 40-w. frosted lights were mounted on *S*'s side of the one-way screen above and to the side of the test-cage.

Procedure: (1) Food response to clicks. The *Ss* were trained initially to retract the lids and retrieve a food-reward. They were then trained to make this response at the presentation of the clicks. This was done by keeping the lids locked until the onset of the clicks. The training was continued until *S* responded rapidly to the clicks, and did not attempt to retract the lids when clicks were not being sounded. There were no solutions presented in the tubes during this period. The *Ss* received 30 trials a day, generally 6 days a week, a schedule that was maintained throughout the experiment.

(2) Single taste-response. A 1.5% solution of NaCl was presented in each tube. The food-reward was now made contingent upon a 'taste'-response, *i.e.* a response in which the tongue or some other part of the mouth clearly made contact with one of the NaCl solutions. This behavior was conditioned by conventional approximation procedures and by increasing the 'distinctiveness' of the solutions. Such a response was immediately reinforced by the clicks, denoting the unlocking of the lids and the availability of the food-reward. The training was continued for at least two sessions until *S* showed a consistent tendency to respond rapidly to the solu-

tions and to retrieve the reward. When a strong positional preference was displayed toward one of the tubes, *S* was required to taste the solution in the opposite tube.

(3) Double taste-response. In the final preliminary step, the food-reward was made contingent upon taste-responses to both NaCl solutions. The behavior was so conditioned that the two responses occurred within 5 sec. of each other. This was done in order to maximize the probability that, in the discrimination proper, *S* would taste both solutions and thus make a simultaneous rather than a successive type of discrimination. It was assumed on the basis of other data that learning would be facilitated and that the thresholds would perhaps be lower under these conditions.[5]

The entire preliminary training required from two to five weeks. It was unnecessary to discard any of the *Ss* during the course of the experiment.

(4) Taste-discrimination. A 1.5% NaCl solution and distilled water were presented separately in the two tubes. Thirty trials a day were given, using a non-correction technique, in which the position of the solutions was randomized in accordance with a predetermined balanced order. The reward was always placed in the foodcup beneath the salt solution. Retraction of a lid defined the end of the discriminative response. For five of the *Ss*, the reward consisted of two 45-mg. glucose pellets. The other three *Ss* received one-quarter of an unsalted peanut as the reward. The training was continued until *S* made 25 or more correct responses on the 30 trials during each of two consecutive test-sessions. The intertrial-time in this procedure was approximately 30 sec.

A dual criterion was used in all the discriminative tests to define an adequate taste-response. The lids were released either 5 sec. after *S* tasted the first solution, or immediately after *S* tasted the second solution, whichever occurred first. Tasting both solutions was not made mandatory since *S* might under some conditions have been prepared to make a choice on the basis of a single taste-response; a second response under these circumstances might have interfered with the discrimination. A record of the pattern of taste-behavior provided information on these questions.

To minimize any use of nontaste cues in the discrimination, the following precautions were observed. The fluid contents in the tubes were varied from session to session. The tubes were frequently resprayed with aluminum paint and were interchanged. Fresh solutions of the salt were used frequently to minimize contamination from the taste-responses. Both the salt and the distilled water were allowed to reach thermal equilibrium with the room that the temperatures of the solutions were the same. The turntable was rotated frequently between trials without changing the position of the tubes, precluding the use of any tactual vibratory or auditory cues.

(5) Threshold-determination. Upon completion of the discriminative tests with the 1.5% NaCl, tests were made with a progressively descending series of logarithmically scaled concentrations of the salt. For any given session, a single concentration was used. The distilled water was continued as the negative stimulus throughout. Each solution of the salt was three-fourths as concentrated as the preceding solution, as follows: 1.125, 0.843, 0.633, 0.475, 0.356, 0.267, 0.200, 0.150, 0.113, 0.084 0.063, 0.048, 0.036, 0.027, 0.020, 0.015, 0.011, each value repre-

[5] K. H. Pribram and Mortimer Mishkin, Simultaneous and successive visual discrimination by monkeys with inferotemporal lesions, *J. comp. physiol. Psychol.*, 48, 1955, 198-202; also, C. P. Richter and A. MacLean, Salt taste thresholds of humans, *Amer. J. Physiol.*, 126, 1939, 1-6.

senting the percentage of salt by weight. When it became apparent that the thresholds were somewhat lower than had been anticipated, alternate steps in the series were skipped below 0.063% during the initial determination. *S* was tested with a given concentration of salt until it made at least 25 correct responses during a single session, or until it failed for a total of three consecutive sessions. If *S* met criterion, the descending series was continued. If *S* failed on three consecutive tests, it was retested at the previously discriminated concentration or subsequently at a higher concentration until it again met criterion. A second descending series of tests was then made starting from this value, but now no steps in the series were skipped. The procedure was continued until *S* failed to meet criterion on a given concentration for a total of six tests including the initial series, or until *S* had failed three times each at some given concentration and at the next higher value which had initially been skipped. The threshold was taken as the lowest concentration above these values at which *S* had met the discriminative criterion.

RESULTS

The results which were obtained on the discriminative tests are summarized in Table I. The *Ss* required on an average 439 trials in order to meet our criterion of discrimination on the 1.5% solution. The differences in performance associated with the two rewards were not statistically significant by a two-tailed Mann-Whitney *U*-test ($p = 0.25$). Indeed, the differences are exaggerated by the fact that one of the *Ss* that received the peanut reward actually made 25 or more correct responses after about half the number of sessions that it required to meet criterion; its day-to-day performance was somewhat erratic. In one animal receiving the glucose reward, the taste-response had to be 'reshaped' after the first 450 trials when it became apparent that *S* was beginning to make pseudo taste-responses; these trials are included in the tabulation of its performance.

TABLE I

TASTE DISCRIMINATION

	Reward				Over-all Mean
	Mean	Range	Mean	Range	
Trials to criterion 1.5% NaCl discrimination*	366	90–660	560	390–660	439
NaCl threshold: % concentration (geometric means)	.0212	.0150–.0356	.0431	.0356–.0475	.0277

* Excludes criterion-trials.

The threshold-values for NaCl, which are also presented in Table I, showed relatively good consistency. An over-all geometric mean thresold of 0.0277% was obtained. Individual values ranged from 0.0150% to 0.0475%. Part of this variability appeared to relate to the type of food-

reward that was used. Thresholds were lower for the glucose than for the peanut reward, a difference that was significant at the 5% level by a two-tailed Mann-Whitney *U*-test.

The pattern of taste-responses was analyzed for both the last 30 trials of the 1.5% NaCl discrimination and the threshold-discrimination. Since changes in the behavior were not apparent for these tests, the data were treated together. An orderly manner of responding to the two discriminanda was shown by the *S*s. Six of the eight *S*s almost always tasted the salt immediately before a correct response. Thus if the water was sampled initially, then *S* would invariably make a second response to the salt before making the food-choice. On the other hand, if the salt was sampled initially, *S* would generally remain the full 5 sec. at the salt solution. Multiple responses from salt to water and back to the salt were much less common. Two of the *S*s which received the peanut reward were deviant in the above patterns of responding. Their food-choices were not consistently preceded by responses to the salt.

Discussion

The results demonstrate clearly the practicability of studying taste-discrimination in animals by employing the taste-cues as discriminative stimuli for a separate food-reward. The reliability of the method is suggested by the finding that the lowest concentration discriminated by *S* on the initial descending series generally corresponded to its threshold following the second set of observations. A subsequent determination, made several weeks after the initial tests, also confirmed the stability of the threshold values.

From a comparative point of view, the results of the present study would place the taste-threshold of the rhesus monkey for NaCl in the same range as that of man and above that of the rat. The lowest thresholds that were obtained in man by Richter and MacLean, for conditions in which the *S* was allowed to sample freely both NaCl and water, averaged 0.016%; this represented the minimal concentration of NaCl which *S* could discriminate from water, and not the concentration (0.087%) at which *S* recognized a salty taste.[6] The minimal values found in the present study for the monkey averaged 0.028%, which is of the same order of magnitude as the Richter and MacLean data. Below this range of NaCl thresholds are those found in the rat for a forced discrimination based on the use of shock to punish incorrect responses. Carr obtained values of

[6] Richter and MacLean, *op. cit.*, 5 f.

0.002 to 0.019%, while those of Harriman and MacLeod were even lower: 0.000025 to 0.002%.[7] The preference-threshold for NaCl in the adrenalectomized rat also approximates Carr's values.[8] More definitive comparisons among the species must, however, await analysis of the various differences in the experimental conditions of the studies mentioned above.

The differences in NaCl threshold associated with the type of food reward must be interpretated with caution. Although a moderately sweet substance such as glucose may indeed increase sensitivity to NaCl, the results obtained may relate to other properties of the rewards, such as their reinforcing value.[9] It is also possible that a sampling bias accounts for the difference in thresholds, since two of the *Ss* that received the peanut reward showed a deviant pattern of taste-behavior. This might have been due to intrinsic individual differences which affected many aspects of their discriminative performance. Support for this supposition comes from a subsequent threshold-test in which the differences obtained for the two rewards were no longer present.

One further implication of the findings on taste-behavior may be noted. The *Ss* appeared to base their discrimination primarily upon the salt solution. That is, they generally sampled the salt just before making a choice. If this pattern related to the gustatory properties of the salt rather than to the food reward's being placed beneath the salt, then some gain in discrimination would seem to have been derived by utilizing a simultaneous rather than a successive type of discrimination, since in the latter the salt solution would have been presented on only 50% of the trials.

SUMMARY

A group of eight food-deprived monkeys was trained to discriminate between a 1.5% NaCl solution and distilled water for a food-reward. Absolute thresholds for NaCl were then determined by progressively decreasing the concentration of the salt. The geometric, mean threshold for discrimination of NaCl was 0.0277% by weight. Individual values ranged from 0.0150 to 0.0475%. The results place the salt-threshold of the rhesus monkey in the same range as that of man, and above that of the rat.

[7] Carr, *op. cit.*, 379; Harriman and MacLeod, *op. cit.*, 468-470; see also S. D. Koh, Absolute taste thresholds of rats obtained by a psychophysical method, Doctoral dissertation, Harvard University, 1958, in which a value of 0.0007 molar concentration, equal to 0.0041% by weight, was found.

[8] Richter, Salt taste thresholds for normal and adrenalectomized rats, *Endocrinology*, 24, 1939, 367-371; J. K. Bare, The specific hunger for sodium chloride in normal and adrenalectomized rats, *J. comp. physiol. Psychol.*, 42, 1949, 242-253.

[9] Pfaffman, *op. cit.*, 1156.

THE IDENTIFICATION OF CONCEPTS AS A FUNCTION OF AMOUNTS OF RELEVANT AND IRRELEVANT INFORMATION

By CLINTON M. WALKER, Hughes Aircraft Company, and LYLE E. BOURNE, JR., University of Utah

Several recent studies have shown that the efficiency of concept-identification depends upon the complexity of the stimulus-patterns to be categorized. Specifically, mean errors to attainment of the concept increased linearly with complexity in these studies.[1] Complexity was quantified in terms of the number of binary dimensions, *e.g.* color (red-green), within which the patterns could vary. Any such dimension was either relevant to solution, *i.e.* necessarily used in classifying the patterns, or irrelevant, hence useless for correct identification of the patterns. As the number of dimensions increased, the number of alternative patterns increased along with the information contained in each alternative. The amount of information (in *bits*) was defined as $\log_2 x$, where $x =$ the number of possible patterns.[2] Since each dimension was binary, each contributed one bit of information.

In previous studies of the effects of complexity, the amount of relevant information was held constant (two bits) while the irrelevant information was varied (one through six bits). Performance, however, should depend upon the amount of information necessarily used in solution as well as upon the amount irrelevant to solution. The purpose of the present experiment was to investigate the relationship between concept-identification and the amount of relevant stimulus-information under varying conditions of irrelevant information.

---

* Received for publication November 30, 1959. This study was supported in part by a grant from the University of Utah Research Fund and in part by a research grant (M-1759) from the National Institute of Mental Health, United States Public Health Service.

[1] E. J. Archer, L. E. Bourne, Jr., and F. G. Brown, Concept identification as a function of irrelevant information and instructions, *J. exp. Psychol.*, 49, 1955, 153-164; Bourne, Effects of delay of information feedback and task complexity on the identification of concepts, *ibid.*, 54, 1957, 201-207; L. E. Bourne, Jr., and R. B. Pendleton, Concept identification as a function of completeness and probability of information feedback, *ibid.*, 56, 1958, 413-420; F. G. Brown and E. J. Archer, Concept identification as a function of task complexity and distribution of practice, *ibid.*, 52, 1956, 316-321.

[2] C. E. Shannan, A mathematical theory of communication, *Bell Syst. tech. J.*, 27, 1948, 379-423; 623-656.

While an increase in irrelevant information increases the complexity of each stimulus-pattern, it has no effect upon the number of categories into which the patterns must be sorted. An increase in relevant information, however, results in an exponential increase in the number of categories, this number being 2^k for k-binary *independent*, relevant dimensions. Bits of stimulus-information, either relevant or irrelevant, might not be independent but rather *redundant* to a certain degree. If the levels of any two or more dimensions are perfectly correlated, the information they contain would be fully redundant and, in the case of relevant information, the categories required by one dimension would be identical for the others. Indications as to how redundant dimensions affect concept-identification have been reported.[3] The present study concerns the effect of variation in independent dimensions only.

Restle has proposed a mathematical theory of discriminative learning in which the stimulus-situation is described as a set of cues, either relevant or irrelevant to a particular response.[4] By assuming (*a*) that relevant cues are conditioned to a correct response, (*b*) that irrelevant cues are neutralized and rendered non-functional in the course of learning, and (*c*) that the probability of a correct response is proportional to the number of cues conditioned to it, he achieved some success in predicting the results of experiments on two-choice discriminative learning. The proportion of relevant cues, θ, was assumed by Restle equal to the rate of learning and was typically estimated from part of the data gathered in any experiment. A parameter, similar to θ, can be obtained directly from the content of the stimulus-series used in a study of concept-identification, if it is assumed that the number of relevant and irrelevant cues is a function of the amount of relevant and irrelevant information in the patterns. With appropriate modification of Restle's Equation [8], the following function is applicable for predicting approximate number of errors to solution under any condition of stimulus-content, assuming independent bits of information:

$$E_t \cong d\left\{\frac{1}{2}+\frac{1}{2}\left[\frac{\log \theta'}{(1-\theta')\log(1-\theta')}\right]\right\}, \qquad [1]$$

where θ' is the ratio of relevant to total stimulus information. For a more complete development of this argument, see Walker.[5]

[3] L. E. Bourne, Jr., and R. C. Haygood, The role of stimulus-redundancy in the identification of concepts, *J. exp. Psychol.*, 58, 1959, 232-238.

[4] Frank Restle, A theory of discrimination learning, *Psychol. Rev.*, 62, 1955, 11-19.

[5] C. M. Walker, Concept identification as a function of relevant and irrelevant information, Unpublished Doctoral dissertation, University of Utah, 1957, 4-18.

A matrix portraying the relationship between stimulus-content and performance, enveloping Equation [1], can be constructed as a model. The *location* of an element within the matrix denotes the degrees of relevant and irrelevant information in the treatment yielding that element. The *value* of the element is the predicted mean error-score for that treatment. Allowing the first digit of each element to represent amount of relevant information and the second amount of irrelevant information in bits, the general model can be schematized as follows:

$$E_t \cong \begin{Bmatrix} 1\text{–}1 & 1\text{–}2 & 1\text{–}3 & \cdots \\ 2\text{–}1 & 2\text{–}2 & 2\text{–}3 & \cdots \\ 3\text{–}1 & 3\text{–}2 & 3\text{–}3 & \cdots \\ \vdots & \vdots & \vdots & \end{Bmatrix} \qquad [2]$$

In this model, the theoretical number of errors to solution, E_t, for a given θ' is based upon an experimentally determined parameter of pattern-discriminability, d, the value of which must be estimated from part of the data.

Method

Task. As in earlier studies of concept-identification, *S* was presented with a series of geometric patterns, each of which was a combination of the levels of *x*-relevant and *y*-irrelevant stimulus-dimensions. To identify the category to which a pattern belonged, *S* responded by pressing one of two, four, or eight buttons, depending upon the number of independent, relevant stimulus-dimensions. The buttons corresponded to the levels (or combinations of levels) of the relevant binary dimension(s). A dimension was relevant to solution if it necessarily was used in correct identification of the patterns. If a dimension was irrelevant, it appeared at each of its two levels within the series of patterns but could not be used to classify the patterns correctly. The criterion of problem-solution was 16 consecutively correct identifications. For a more complete discussion of the task, see Bourne.[6]

Design. A $3 \times 3 \times 2$ factorial design was used with three levels of relevant information (1, 2, and 3 relevant dimensions), three levels of irrelevant information (1, 2, and 3 irrelevant dimensions), and two different problems (defined in terms of the actual dimensions used as relevant and irrelevant).

Stimulus-lists. Both the relevant and the irrelevant dimensions were selected from a population of six simple binary attributes, viz. color (red or green), number (one or two figures), size (large or small), form (square or triangle), orientation (upright or tilted), and horizontal position (on the left or right). Since there were two levels on each dimension, the amount of information in the stimulus-objects could be quantified simply. Since the number of alternative stimulus-patterns doubled with the addition of each dimension, evaluation of $\log_2 (2^N)$, where N is the number of

[6] Bourne, *op. cit.*, 201.

dimensions, shows that the number of bits of information is equivalent to the number of stimulus-dimensions present. In the problems studied, then, there were 1, 2, or 3 bits of relevant and 1, 2, or 3 bits of irrelevant information. The number of response categories was equal to the number of alternative relevant stimulus-combinations. Therefore, one, two, and three relevant dimensions required, respectively, the learning of two-, four- and eight-categories.

Eighteen series of patterns were prepared on loops of white-vellum tape to present problems at the nine relevant-irrelevant conditions with two sets of relevant and irrelevant dimensions. Nine lists used size, size and number, or size, number, and horizontal position to display relevant information. The other nine lists used color, color and form, or color, form, and orientation as relevant. The irrelevant information was displayed by the color-group when the size-group was relevant and vice versa.

Apparatus. The stimulus-patterns were presented to *S* at a rate of one every 6 sec. on a Gerbrands memory-drum. *S* sat at a distance of about 2 ft. from the drum with the response-board before him. The buttons on *S*'s board activated lights on a control-panel located behind a partition and out of *S*'s view. Thus *E* could record each response and in turn inform *S* as to the correctness of each response. Only during the first 3 sec. of the 6-sec. trial did the stimulus-pattern remain in the window of the drum. The movement of the drum to a blank position was *S*'s signal to respond. Information was given by a light-panel placed directly above and about 4 in. from the keys on the response-unit. The panel consisted of a row of eight jewelled lamps, one directly above each of the eight response-keys. Regardless of *S*'s response, the lamp above the correct key for any given pattern was lighted during the last second of the trial. A covering for certain of the lamps and keys was used when the number of categories was less than eight, *i.e.* with one and two relevant dimensions.

Subjects. The *S*s (162 in number) were students in courses in elementary psychology and were assigned in order of appearance to solve one of the problems in one of the nine relevant-irrelevant combinations. Every *S* was presented at the outset with detailed instructions as to the nature of the task, the operation of his controls, the meaning of the lights, and the criterion of problem-solution.

Results and Discussion

Two response-measures were used: trials to solution and number of errors. Due to the high correlation between these measures ($r = 0.91$) and the fact that instructions to *S* stressed accuracy in problem-solving, number of errors was considered the main dependent variable.

The following is a 3×3 matrix which portrays the obtained errors to solution, E_o, under the various conditions of relevant and irrelevant information:

$$E_o = \begin{pmatrix} 5.3 & 8.6 & 13.0 \\ 21.2 & 30.5 & 32.5 \\ 78.7 & 131.8 & 155.1 \end{pmatrix}.$$

Columns represent 1, 2, and 3 bits of irrelevant information and rows 1, 2, and 3 bits of relevant information (two-, four-, and eight-choice learning, respectively). An analysis of variance was performed on these data. Two main effects, amount of relevant and irrelevant information, were significant at the 1% level ($F = 118.66$ and 8.60, respectively, $df. = 2$ and 144). The significance of the amount of irrelevant information is consistent with all earlier studies and indicates that the number of errors made in solving this type of problem is an increasing function of the

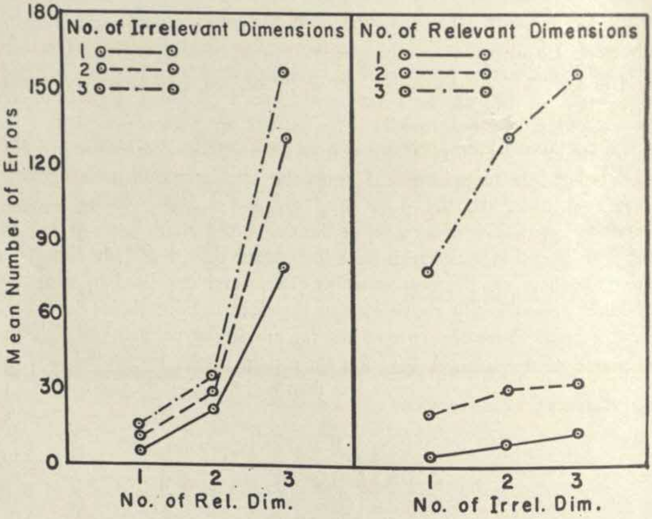


FIG. 1. MEAN ERRORS TO SOLUTION AS A FUNCTION OF AMOUNT OF RELEVANT AND IRRELEVANT INFORMATION
(Every point represents the data from 18 *Ss*.)

complexity of the pattern. Fig. 1 shows that this increase nearly follows a straight line. In an orthogonal polynomial analysis applied to the relationship, only the linear term reached the 1% level of significance.

Amount of relevant information produced a greater effect on performance. As can be seen in Fig. 1, errors increased at a positively accelerated rate with relevant information. An orthogonal polynomial analysis of this curve statistically confirmed its curvilinearity; the quadratic component reached significance at the 1% level.

Somewhat inconsistent with earlier results was the failure of the variance identified with problems to reach significance ($F = 3.48$, 1 and 144 $df.$, $P > 0.05$). The effect of problems has, however, always been small

and dependent on the discriminability between the levels of the dimensions used as relevant.

The effect of amount of relevant information depended on the level of irrelevant information employed in a problem. This is evidenced by the lack of parallelism among the sets of curves in Fig. 1 and by the significance of the interaction of relevant by irrelevant information ($F = 7.72$, $df. = 4$ and 144, $P < 0.01$). This interaction itself was somewhat dependent upon which of the two sets of problems was used, *i.e.* the interaction of relevant information, irrelevant information, and problems was significant ($F = 6.06$, $df. = 4$ and 144, $P < 0.01$).

With trials to solution as the measure, essentially identical results were found; the analysis of variance showed the same sources of variance to be significant.

The matrix of obtained errors can be reduced by extraction of the overall trend when the amounts of relevant and irrelevant information are equal. The result is the product of a diagonal and a residual matrix:

$$E_o = \text{diag.}\,(5.3, 30.5, 155.1)\begin{pmatrix} 1.0 & 1.6 & 2.4 \\ 0.7 & 1.0 & 1.1 \\ 0.5 & 0.9 & 1.0 \end{pmatrix}.$$

For comparison, computations using Equation [1] yield the following theoretical array:

$$E_t \cong 1.5d\begin{pmatrix} 1.0 & 1.7 & 2.5 \\ 0.7 & 1.0 & 1.3 \\ 0.6 & 0.8 & 1.0 \end{pmatrix}.$$

Examination of the diagonal multiplier-matrix from the data indicates that its elements form a power sequence. The scalar multiplier of the model, $1.5d$, should be replaced by the sequence, $(1.5d)^n$, where n is the number of (equal) relevant and irrelevant dimensions, in order to describe the data adequately. Using the obtained errors from the condition with one relevant and one irrelevant bit of information, $d = 3.53$. The theoretical diagonal matrix is then diag. (5.3, 28.1, 148.9). The closeness of this fit should allow the tentative formulation of a simple equation for predicting errors to solution in a concept-identification problem with equal relevant and irrelevant information:

$$E_t \cong (1.5d)^n.$$

It is perhaps worth noting that a simple transformation, $x' = 1 / x$, appled to the elements, x, above the main diagonal in the residual error

matrix reduces the range of those elements to that of the elements below the diagonal, *i.e.* 0 to 1. But, more interesting, symmetric matrices result.

$$E_t \cong \text{diag.}\left[(1.5d), (1.5d)^2, (1.5d)^3\right]\begin{pmatrix} 1.0 & 0.7 & 0.4 \\ 0.7 & 1.0 & 0.8 \\ 0.6 & 0.8 & 1.0 \end{pmatrix}, \quad \text{and}$$

$$E_o = \text{diag.}\,(5.3, 30.5, 155.1)\begin{pmatrix} 1.0 & 0.7 & 0.4 \\ 0.7 & 1.0 & 0.9 \\ 0.5 & 0.9 & 1.0 \end{pmatrix}$$

The results of this experiment clearly support several hypotheses generated from the proposed model. As predicted, and as has been shown in earlier studies, the difficulty of the task was affected by variation in the amount of irrelevant information in the patterns. Performance was an inverse linear function of irrelevant information. Amount of relevant information, however, had a much greater effect on difficulty. As the amount of information necessarily used to classify the patterns increased, performance decreased exponentially. Close fits of the predictions to the data clearly support the assumption that θ, the proportion of relevant stimulus-cues, is functionally related to the proportion of relevant stimulus-information. Bricker found results similar to these with an experiment designed to investigate the effect of both independent and redundant relevant information in a different task.[7] It is important to note that relevant information, as varied in the present experiment, is information that *S* must use to classify any given pattern. An experiment reported earlier, investigated concept-identification as a function of redundant, relevant stimulus-information.[8] Each bit of relevant information in the sequence of stimulus-patterns was redundant with the first, *i.e.* added no new stimulus-information necessarily used in the correct classificatory scheme. As might be expected in this case, increases in relevant information enhanced performance since *S* could use any one of the relevant bits to categorize correctly.

SUMMARY

Theoretical formulations of Restle coupled with modifications suggested by information-theory were used to construct a matrix-model for concept-identification.[9] Simply stated, this model holds that performance, measured

[7] P. D. Bricker, The identification of redundant stimulus-patterns, *J. exp. Psychol.*, 49, 1955, 73-81.
[8] Bourne and Haygood, *op. cit.*, 232.
[9] Restle, *op. cit.*, 15.

by total errors to concept-attainment, is a function of stimulus-discriminability and number of relevant and irrelevant bits of information in the stimulus-patterns.

An experiment, in which the independent variables were relevant and irrelevant information, was conducted as a test of the model. A 3 × 3 × 2 factorial design, with three levels of relevant information, 1, 2, and 3 bits, three levels of irrelevent information, 1, 2, and 3 bits, and two different problems, was used. Each of 162 *S*s served individually and learned to categorize visually presented geometric patterns.

The model predicted, and the results confirmed, a linear decrement in performance with increased irrelevant information. This finding corroborated earlier experimental evidence. In addition, an exponential performance-decrement resulted from increased relevant information as predicted.

FORMULATION AND REFORMULATION OF FIGURE-CONCEPTS

By DONALD M. JOHNSON, Michigan State University

The research reported here is based on an analysis of problem-solving into two sequential processes, designated as *preparation* and *solution*. The approach adopted has yielded reasonable results when applied to a variety of short problems and to comparisons between verbal analogies which emphasize induction and those which emphasize deduction.[1] The relation between time spent on preparation and the number of items presented has been determined,[2] as well as the relative effects of relevant and irrelevant items.[3] The present experiment concerns problems in which *S* has (1) to discover the common property of a group of figures and then (2) to identify another instance of the concept. It is assumed that the preparatory activity in such cases consists of an initial formulation of the problem. *S* surveys the given material, observing some properties while ignoring others, and achieves a tentative formulation or model to guide his next step. Solution consists of a search for a figure that matches this formulation. It follows from this assumption that *S* will not search for a figure with properties that he has ignored in his formulation. It follows also that if *S* does not find a solution that matches his formulation, he will reformulate the problem and search for another type of solution.

To test this simple model of cognitive dynamics, problems are needed that permit two distinct formulations and two distinct solutions from which the antecedent formulations can be inferred, as well as an independent variable that will slant the formulation toward one or the other. The problem illustrated by Fig. 1 is suitable for this purpose, because the figures on the left can be conceptualized in terms of shape or texture. With such a card in an exposure apparatus, the instructions are to find what the figures on the left have in common and then to find on the right an-

* Received for publication February 1, 1960. This research was supported by a grant from the National Science Foundation.

[1] D. M. Johnson, Serial analysis of thinking, *Ann. N. Y. Acad. Sciences*, 91, 1960, 66-75.

[2] D. M. Johnson, R. E. Lincoln, and E. R. Hall, Amount of material and time of preparation for solving problems, *J. Psychol.*, 51, 1961, 457-471.

[3] Johnson and Hall, Organization of relevant and irrelevant words in the solution of verbal problems, *J. Psychol.*, 52, 1961, 99-104.

other example of this class of figures. Fig. 1 is an ambiguous problem in that it can be formulated and solved in terms of the shape of the figures or in terms of their texture. It is a safe assumption that an *S* who selects No. 1 has formulated the problem in terms of shape, and that one who selects No. 10 has formulated it in terms of texture.

Rees and Israel, Luchins, and others have demonstrated that a set for a certain type of solution can be induced by training on a short series of problems for which such solutions are obvious,[4] hence two kinds of pre-test problems were used to slant the *S*s toward either shape or texture. The first hypothesis states that *S*s who work on a series of problems that are

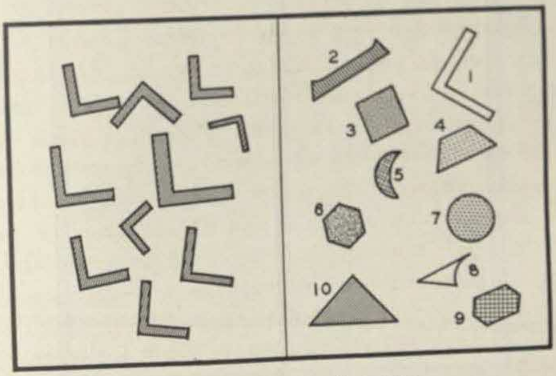


FIG. 1. AMBIGUOUS PROBLEM

easily solved in terms of only one common property will establish a set to formulate ambiguous problems in this way and select a solution that matches this formulation.

The problems used to study reformulation are similar except that there is only one possible solution on the right side, and these problems are presented to *S*s who are set to search for solutions of the other class. The second hypothesis states that when *S* does not find a solution in accord with the training series he will switch back to the preparatory material on the left side, reformulate it in terms of the other common property, return to the right side, and select the correct solution. The exposure-box permits *S* to switch back to the preparatory side, and such a switchback is taken as an objective sign of reformulation.

[4] H. J. Rees and H. E. Israel, An investigation of the establishment and operation of mental sets, *Psychol. Monogr.*, 46, 1935 (No. 210), 1-26; A. S. Luchins, Mechanization in problem solving: The effect of *Einstellung*, *ibid.*, 54, 1942, (No. 248), 1-95.

Procedure

The serial-exposure box, shown in Fig. 2, consists of two adjacent chambers, separately lighted. The front of the box is a half-silvered mirror, and at the back is a holder for 5 × 8 cards so situated that half of a card is visible in each chamber. The left half is exposed to *S*, and then, when he is ready, he throws a centrally mounted switch that shifts the light to the right chamber, exposing the right half of the card. When he has selected a solution, he pushes an appropriately labelled button. Certain groups were permitted to switch back to the left side after attempting solution, and for such groups one switchback was forced during instructions. The time spent, after switchback, on the second preparation was automatically cumulated with the time spent on the initial preparation. Thus, for each problem, a record was obtained of time spent on preparation, time spent on solution, the solution selected, and, for certain groups, the number of switchbacks.

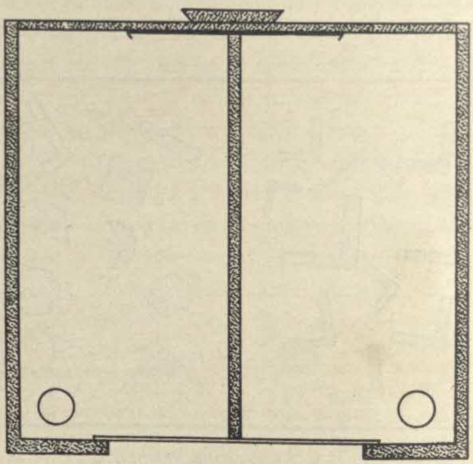


FIG. 2. FLOOR PLAN OF EXPOSURE BOX

The problem, printed in two parts on a 5 × 8 card, is inserted in the holder on the back. Visibility of each part through the half-silvered mirror is controlled separately by the light at each side.

Two blocks of eight easy problems each were constructed for the two training series, and half of each group did one or the other of these before attempting the final problems. One block was composed of problems that could be solved only in terms of shape and the other only in terms of texture. These were shuffled before presentation to *S*.

Four groups of 24 *S*s each, recruited from introductory psychology classes, were required to solve three problems. These problems were presented immediately after the eight training problems, in fixed order, rotated among *S*s to balance order-effects.

Ambiguous problems. One group did three ambiguous problems, similar to Fig. 1, with the hypothesis that the subgroup trained on shape would formulate the problems in terms of shape and select a solution of the same shape, while the sub-

group trained on texture would select a solution in terms of texture. If this hypothesis is confirmed, we can assume that the two training series do determine the formulation, and we will have a method for the study of reformulation.

Reformulation. Similar problems were used to test for reformulation, except that there was only one solution for each. The subgroup trained on shape was given three problems soluble only in terms of texture, and the subgroup trained on texture was given three problems soluble only in terms of shape. It was expected that these problems would be more difficult than the ambiguous problems in respect to time and errors, and that they would result in more switchbacks.

Control I. Since the problems used for reformulation had only one correct solution, they might be considered more difficult than the ambiguous problems which had two correct solutions. To check on difficulty independent of reformulation, these problems of only one solution were given to another group, both subgroups of which had training problems in accord with the test-problems.

Control II. It is implicit in this procedure that, when a series of problems is preceded by problems solved in a different way, the *S* who does not reformulate

TABLE I

ANALYTICAL DATA ON FOUR GROUPS SOLVING FIGURE-CONCEPT PROBLEMS

Group	Preparation-time	Solution-time	Set solutions	Errors	Switch-backs
Ambiguous	3.9	3.7	64	0	9
Reformulation	6.1	7.4	—	12	43
Control I	4.4	4.0	70	2	9
Control II	6.7	11.3	—	31	—

the problems will not be able to solve them, or at least he will make more errors. To check on this implication, another group was given the same test-problems and the same training problems as for reformulation, but with switchback prohibited by the instructions.

RESULTS

Ambiguous problems. The 12 *Ss* of the subgroup trained on shape selected solutions on the basis of shape for all 36 problems. The 12 *Ss* trained on texture selected solutions on the basis of texture for 28 of the 36 problems. (The other 8 solutions were based on shape.) Since there were 10 possible solutions displayed, the probability of choosing any one is 0.1, and since each *S* solved three problems, a chance score is 0.3. The obtained scores ranged from 0-3 with a mean of 2.3, which is significantly different from 0.3 at the 1% level. For the whole group of 24, there were 64 out of 72 solutions consistent with the set presumably established by the training series—called 'set' solutions in Table I.

Table I also gives the geometric means, in seconds, for preparation-time and solution-time. Fig. 3 shows that there was little change in the times from the first of these ambiguous problems to the third.

It is clear that this procedure does establish a strong set to formulate

the three problems in accordance with immediately preceding experience. Thus, one of the determinants of formulation has been demonstrated, and a procedure is available for testing the hypothesis about reformulation.

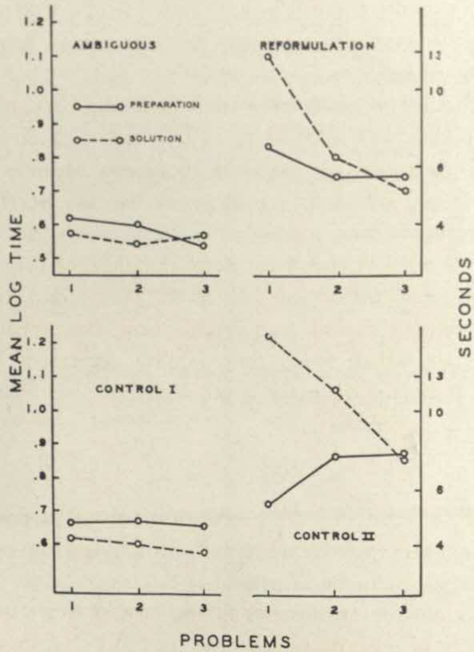


FIG. 3. TIME SPENT ON PREPARATION AND ON SOLUTION OF THREE PROBLEMS BY ORDER OF PRESENTATION

Each point is the mean of 24 times.

Reformulation. As expected, every *S* switched back to the preparatory material at least once, and the total number of switchbacks was 43, to be compared with 9 for the ambiguous problems. Giving each *S* a score for number of switchbacks, the mean score for the ambiguous group is 0.4, and for reformulation 1.8, a difference which is highly significant. Disregarding the number of switchbacks and considering only the number of *S*s who switched back at least once, the comparison is between 6 and 24, or between 25% and 100%, also highly significant. It is clear then that when *S* is faced with a problem for which his formulation, as defined by this procedure, is inadequate, he is very likely to have another look at the preparatory material and to reformulate the problems.

Fig. 3 (Reformulation) shows that preparation-time, which includes

time of switchback and reformulation, was long for the first problem and decreased rapidly. The number of switchbacks likewise dropped from 24 on the first problem, to 12 on the second, and 7 on the third. These results also are consonant with the assumptions about reformulation.

Control I. The results for this group, shown in Table I and Fig. 3, are quite similar to the results for the ambiguous problems, hence the conclusions need not be altered.

Control II. This group, prohibited from switching back, made 31 errors, while the group that was allowed to switch back and presumably to reformulate made 12 errors on the same problems. Mean errors were 1.3 and 0.5, a difference which is significant at the 2% level. Considering only the number of *Ss* who made one or more errors, the comparison is between 18 for Control II and 8 for Reformulation, or between 75% and 33%, which is significant at the 1% level. This group, like the group allowed to reformulate, spent considerable time on preparation for the first problem, after which mean time rapidly decreased. In general, if reformulation (as herein defined) is required but switchback is not permitted, time and errors increase.

Discussion

As a method for studying the processes involved in problem-solving, serial exposure appears to work quite well. Switching from one side of the card to the other seems to be as natural as the concomitant movements of head and eye. Time for adjustment to the task is negligible, and there are no complaints of interruption of the continuity of thought. The results presented here could not have been obtained without assuming some temporal differentiation within the problem-solving episode, even within an episode of the order of 15 sec. The curves for Reformulation and Control II (Fig. 3) show that the separate timing of preparation and solution yields results that could not have been obtained by over-all timing. Previous research has demonstrated that serial exposure is no less efficient than simultaneous exposure in respect to time and errors.[5]

The results at hand can be explained by the assumption that the first or preparatory step in the solution of these problems is some kind of formulation of the data in view, and that this formulation controls the next step, the search for a solution. The formulation may be said to have a facilitative effect since, when one of the available solutions fits the formulation, solution occurs in about 4 sec. The formulation also has a negative

[5] Johnson, Lincoln, and Hall, *op. cit.*, *J. Psychol.*, 51, 1961, 457-471.

effect, since data not included in the formulation cannot be utilized in the search for a solution.

The formulation can be called a 'cognitive pattern' since it represents knowledge about the perceived figures. It also can be called an 'instrumental pattern' or 'search-model' since it is constructed in order to solve a problem. From the investigator's point of view, such a pattern can be treated as a 'hypothetical construct' or 'mediating variable' since it is known only by inference from its antecedents and consequents. The above results show that it is possible to manipulate antecedent conditions of formulation and make reasonable predictions about consequents. In fact, when a series of problems is arranged to force reformulation, and switchback is allowed, it is possible to identify and time, for all *S*s on all problems, four operations in sequence: (1) the first formulation of the problem; (2) unsuccessful search for a solution to match this formulation; (3) switchback and reformulation; and (4) successful search for a solution to match the second formulation.

SUMMARY

This research began with a hypothetical description of the solution of problems in terms of two processes, called 'preparation' and 'solution.' In the case of concept-formation problems, presented serially in two parts, preparation was assumed to consist of the formulation of a common property of the ten figures, and solution was assumed to consist of a search for a solution, in a display of ten possibilities, that matched this formulation. Serial exposure under *S*'s control permitted the timing of each operation.

Ambiguous problems, the common property of which was either a shape or a texture, were solved, as predicted, in accordance with the formulation established during solution of a brief training series of easy problems.

When a problem is formulated and no solution is available to match the formulation, reformulation was predicted, the objective sign of which was taken to be switching back to the preparatory material. A series of problems arranged to force reformulation yielded switchbacks as expected. Hence it is claimed that formulation and reformulation have been identified and timed, as have attempts at solution.

LEARNING WITHOUT AWARENESS AND MEDIATED GENERALIZATION

By MELVIN MANIS, VA Hospital, Ann Arbor, and
EDWARD J. BARNES, Michigan State University

This study was designed to evaluate the role of verbal awareness in secondary (or mediated) generalization. The basic plan of the experiment was to identify *Ss* who had learned a discriminative task, but could not verbalize the basis for their responses, and then test them for generalization, using stimuli that were *conceptually related* to those of the learning-series. The two major goals were (1) to determine whether *Ss* who had learned without awareness would show significant mediated generalization, and (2) to see if aware and unaware *Ss* would differ in the amount of generalization that they exhibited. It was anticipated that in learning the original discrimination, the aware *Ss* would come to associate their overt responses with the conceptual meanings of the stimuli, while the unaware *Ss* might rely more heavily on the physical properties of the stimuli (hence their inability to verbalize the principle of the discrimination); it was, therefore, predicted that the aware *Ss* would show more mediated generalization than would the unaware *Ss*.

Of subsidiary interest to the questions regarding generalization was a comparison of the *Ss* with and without insight in terms of speed of learning, *IQ*, and retention. While no prediction was made for the measure of learning, it was anticipated that the *Ss* with insight would show higher *IQ*s and better retention than those without insight.

METHOD

Subjects. The data reported below were obtained from 23 medical and surgical patients in a VA Hospital; most of these men were in their twenties or thirties.

Learning. The *Ss* were asked to play the role of 'plane spotters'; they were shown a series of 'airplane insignia' and were instructed to 'guess' which insignia came from 'friendly planes' and which from 'enemy planes.' After each response, *Ss* were told whether they were 'right' or 'wrong.'

Each insigne consisted of the Arabic numbers '2,' '3,' '4,'' and '5,' arranged in a cross-shaped pattern. Although these same numbers were presented throughout, they appeared in different positions from one insigne to the next, with the restriction that the number '2' (for friendly planes) or '5' (for enemy planes) always

* Received for publication March 7, 1960. Our thanks are due Drs. E. L. Kelly, L. J. Bookbinder, and D. W. Bostian for their assistance on this study.

appeared on the left; there were three distinct (*i.e.* non-identical) insignia of each type.

The stimulus-objects were presented by means of a translucent screen and a Viewlex slide-projector. *E* was seated behind the screen, out of sight of the *S*s. The objects were presented in four different orders, to prevent serial learning; the data were recorded in blocks of 11 trials each and learning-trials were continued until *S* had completed one errorless block. Three *S*s were eliminated because of their failure to reach this criterion within 18 blocks.

Generalization. After the criterion of learning had been achieved, as a test of generalization, *S*s attempted to identify a new set of 20 insignia, in which the Arabic numerals were replaced by different numbers of circles (*e.g.* *two* circles in place of the number '2,' *three* circles for '3,' etc.). This substitution was fully explained and the *S*s were instructed to follow the same scheme (*i.e.* hypothesis) that they had used in the preceding series of trials. They were not given knowledge of results during the test for generalization.

Apart from the division into friendly and enemy insignia, the stimuli which were presented in the test for generalization could be divided into a set of 'old' insignia and a set of 'new' ones. Each of the 'old' insignia was totally equivalent to one of the slides in the learning-series; *i.e.* each numerical element was replaced by its 'circle-equivalent,' without any alteration in position. The 'new' insignia did not have this total equivalence, although all 'new' insignia did have either two or five circles in the left position, and hence could be correctly classified by applying the principle which governed the learning-series. Four 'old' stimulus-objects and six 'new' ones were presented in the test for generalization; these objects were randomly intermingled and were presented twice, in different orders. Thus, the test consisted of 8 trials on the 'old' insignia and 12 on the 'new' ones.

Retention. After being tested for generalization, each *S* was given a set of 22 additional trials on the original stimulus-objects, *i.e.* the 'numbered,' with no feedback, to assess retention.

Insight. When the above procedures had been completed, each *S* was asked the following question: "How did you know which planes were friendly and which were enemy?" If *S* could not verbalize the basis for the discrimination, *E* probed further, until he felt confident that *S*'s lack of insight could not be attributed to limitations in verbal facility. All *S*s who at this point were still unable to verbalize the principle, were then asked: "Take this last slide; you said it was a 'friendly plane'; what made you think it was friendly?"[1]

Although an alternative procedure might have been to check for insight as soon as the learning-criterion had been achieved, this possibility was rejected, since the *S*s might then have felt constrained to follow their stated 'principles' during the test for generalization; this would have the effect of reducing the possibility that *S*s whose verbalizations were *incorrect* would show generalization above-chance. To guard against the possibility that a principle might be formulated during the learning-trials and then rejected as inapplicable, following exposure to the 'changed' stimuli of the generalization-test, the similarity, *i.e.* conceptual equivalence, which existed between the 'number' and the 'circle' objects was fully explained (see above); moreover, *S*s were instructed to continue using whatever

[1] All *S*s correctly identified this stimulus-object as a 'friendly plane.'

scheme they had developed during the learning-series. Examination of the protocols of the *Ss*' interviews suggests that these precautions were successful.

To determine whether *S* belonged in the 'insightful' or the 'not-insightful' Group, the following criterion was used: if *S*'s verbalization was accurate, *i.e.* if application of his stated principle would have led to a correct identification of each object, he was judged to have insight. If not, he was placed in the group without insight. As defined by this criterion, there were 13 *Ss* with insight and 10 without insight. Four *Ss*, whose classifications were doubtful, were removed from the study. Some of the verbalizations given by *Ss* without insight were: "You have to keep track of the numbers. When the five and four are together, it's friendly." "You have to take all the numbers into consideration." "They have to balance to be friendly. The lopsided ones are not friendly. The numbers usually run up and down and across."

Intelligence. Following the procedures described above (which generally required about 30 min.), the *Ss* were given a short form of the Otis Intelligence Test.[2]

RESULTS

Inspection of the data revealed that the *Ss* without insight had performed at a level above chance in responding to both the 'new' and the 'old' stimulus-objects.[3] On the 'new,' 9 of the 10 *Ss* without insight gave performances better than chance, and one performed at a level below chance. These results, using a one-tailed sign-test, are significant beyond the 2% level. On the 'old' stimulus-objects, 8 of the 10 *Ss* without insight gave performances better than chance, one performed at a chance level, and one below chance. These results are significant beyond the 3% level.

Examination of the 'insightless' *Ss*' responses to the interview indicated that some of them had verbalized hypotheses that were correlated (although imperfectly) with the correct principle. Although the performance of the 'insightless' *Ss* on tests for generalization was better than chance, it was thus possible that it was not better than the performance that would be expected on the basis of their partial insights. To control for this possibility, the obtained performance of every *S* without insight on the test for generalization was compared to the level of performance that he would have achieved had he responded in terms of the principle that he verbalized in the interview. On slides that could not be categorized in terms of this principle, it was assumed that *S* would have been correct 50% of the time by guessing. The results clearly indicated that partial insight was *not* an adequate explanation for the results obtained. For the total generaliza-

---

[2] A. S. Otis, *Otis Self-Administering Tests of Mental Ability: Higher Examination; Form B*, 1950.

[3] It was assumed that, on the basis of chance-guesses, the *Ss* would have been correct on 50% of the trials for generalization.

TABLE I

COMPARISON OF *S*S WITH AND WITHOUT INSIGHT ON SIX VARIABLES

Variable	Insightful ($N=13$)		Insightless ($N=10$)		p of diff.
	Mean	*SD*	Mean	*SD*	
Learning (blocks to criterion)	4.2	3.6	8.6	5.2	$<.025$
No. correct: 'new' objects	10.9	2.4	8.0	6.9	$<.0025$
No. correct: 'old' objects	7.8	.4	6.2	2.6	$<.025$
No. correct: old and new	18.7	1.9	14.2	4.8	$<.005$
No. correct: retention	21.2	2.5	18.3	3.5	$<.025$
**IQ*	92.0	12.5	84.6	12.8	$<.10$

* *IQ*-measures were unavailable for 2 *S*s; the means reported, therefore, represent *N*s of 12 and 9 for the insightful and insightless groups, respectively.

tion-series, 9 of the 10 *S*s without insight exceeded their expected levels of performance. This result, as evaluated by a one-tailed sign-test, is significant beyond the 2% level. On the 'new' stimulus-objects, 9 of the 10 *S*s did better than predicted, while on the 'old' objects, 2 *S*s equalled their predicted levels, and 7 of the remaining 8 exceeded them. These results are significant beyond the 2% and 4% levels, respectively.

With vs. without insight. The *S*s with and without insight were compared on six dependent variables: *IQ*, speed of learning, generalization to the 'new' stimulus-objects, generalization to the 'old,' generalization to the 'new' and the 'old' combined, and retention. Table I presents the means and standard deviations of the groups for each of these variables, together with the significance of the differences, as estimated by *t*-tests.[4] Apart from the measure of the speed of learning, where no prediction was made, one-tailed *p*s are reported.[5]

As may be seen in Table I, the insightful *S*s performed more adequately on each of the dependent variables; on the measure of *IQ*, however, the results were barely significant at the 10% level.[6]

(1) Generalization. Although the insightful *S*s showed superior generalization, it seemed possible that this might be attributable to superior associative strength, for

[4] Similar results were obtained when the data were evaluated by means of the nonparametric *U*-test; *IQ* was the only dependent variable which failed to achieve a $p < 0.05$.

[5] Throughout this report, two-tailed *p*s are presented for comparisons of learning speeds, and one-tailed *p*s for comparisons on the other dependent variables.

[6] Preliminary work with several graduate students, most of whom showed a rapid acquisition of insight, suggests that the failure to find a more significant relationship between *IQ* and insight may be attributable to the restricted range of intelligence (65-106) among the experimental *S*s.

previous research suggests that the achievement of a fixed criterion does not imply constant strength of association.[7] To take account of this possibility, a series of analyses of ovariance was carried out, using retention as a control-variable. It was assumed that, at the time of the test for generalization, individual differences in the strength of the original learning would be directly related to differences in retention (as measured immediately afterwards). The results of the analyses of covariance indicated that for both the 'new' and the 'old' stimulus-objects, *Ss* with insight showed better performance than those without insight, even after the effects attributable to associative strength (as estimated by retention) had been eliminated.[8] For the 'new' stimulus-objects $p < 0.013$; for the 'old' objects, $p < 0.07$. For the combined set of objects, the difference between the groups was significant beyond the 2.5% level. When speed of learning was statistically controlled by means of covariance, the differences in generalization were again essentially unchanged, although the level of significance for the 'old' objects was raised to 7%.[9] A final series of analyses of covariance indicated that the differences in generalization were not attributable to *IQ*.[10]

(2) Learning. Although the insightful *Ss* reached the criterion of learning more quickly than did the insightless *Ss* ($p < 0.025$), these results were reduced to the 11% level when *IQ* was controlled through an analysis of covariance.[11]

(3) Retention. As shown in Table I, the *Ss* with insight showed faster learning and better retention than those without insight. Since differences in speed of learning may produce differences in associative strength, and the present procedures did not provide a direct measure of associative strength at the termination of learning, it is difficult to draw any firm conclusions concerning the effects of insight upon retention.

Discussion

The experimental results indicate that in responding to a task requiring mediated generalization, *Ss* without insight may exhibit levels of performance that significantly exceed their ability to verbalize the basis for their responses. This finding is in accord with data reported by Lacey, Smith, and Green;[12] *Ss* in their experiment were instructed to give continuous free-associations to a series of stimulus-words until *E* told them to stop (after 15 sec.), and to tap meanwhile at a constant rate on a

---

[7] B. J. Underwood, Speed of learning and amount retained: A consideration of methodology, *Psychol. Bull.*, 51, 1954, 276-282.

[8] The correlations between retention and generalization were 0.39, 0.29, and 0.38 for the 'new,' the 'old,' and the combined sets of stimulus-objects, respectively. In these analyses of covariance and those reported below, the data conformed reasonably well to the assumptions underlying this procedure; see E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 323-333.

[9] Speed of learning and performance on the test for generalization correlated —0.52, —0.30, and —0.47, for the 'new,' 'old,' and combined sets of stimulus-objects.

[10] *IQ* correlated 0.37, 0.38, and 0.40 with the three measures of generalization.

[11] There was a correlation of —0.25 between speed of learning and *IQ*.

[12] J. I. Lacey, R. L. Smith, and Arnold Green, Use of conditioned autonomic responses in the study of anxiety, *Psychosom. Med.*, 17, 1955, 208-217.

telegraph-key. Half the *S*s were shocked 0.5 sec. after they had been told to stop responding to the word *cow,* which was presented six times. Presentation of this word eventually elicited an increase in heart rate, and most importantly for the present study, this reaction generalized to other 'rural' stimulus-words, even when *S*s were unaware that the word 'cow' had always preceded the shocks.

These results suggest that in a discriminative task involving meaningful stimulus-materials, the *S* who learns without awareness is probably responding to the 'conceptual meaning' of the materials, much as the insightful *S* is. Thus, in the present experiment, *S*s without insight learned to respond to the *numerical concepts* that appeared in the crucial left position, rather than responding solely to the *physical configurations* of the Arabic numbers that comprised the original stimulus-objects. This mode of response enabled them to transfer their learning to the insignia which were presented in the test for generalization, where the concepts that had been shown initially were represented by stimulus-objects which were physically *unlike* those in the original series.

The significant amount of generalization shown by the *S*s without insight supports Leeper's contention that *S*s in tasks involving the formation of concepts often "develop the ability to name new examples without being able to say how they do it, even when the necessary formulations lie well within the limits of their vocabularies."[13] Adams, however, has noted that although *S*s in each of the studies cited by Leeper may have produced verbalizations that differed from those with which the *E* started, it was quite possible that the verbalizations they gave would have provided an adequate basis for the performance they exhibited.[14] In the present study, this possibility was put to a direct test; the results indicated that the amount of generalization shown by the *S*s without insight *exceeded* the level of performance that would have resulted if they had consistently followed their respective statements of the principle of discrimination.

Although it is clear that the mediated generalization can occur in the absence of verbal insight, verbal mediators apparently play a significant role in determining the magnitude of the effect. In the present study, *S*s with insight showed *more* mediated generalization than did those without insight; thus, verbal insight apparently *strengthened* the mediated similarity between the stimulus-objects of the learning series and the associated

[13] Robert Leeper, Cognitive processes, in S. S. Stevens (ed.) *Handbook of Experimental Psychology*, 1951, 731.

[14] J. K. Adams, Laboratory studies of behavior without awareness, *Psychol. Bull.*, 54, 1957, 383-405.

objects in the test for generalization. This result may have come about in the following manner: an *S* with insight would probably implicitly 'label' each stimulus-object in the generalization-series as having either a 'five-on-the-left,' or a 'two-on-the-left'; his *overt* response (*friend* or *enemy*) would then depend upon which of these mediating responses had been elicited. If this pattern were applied consistently during the test for generalization, performance would be perfect, since *S*'s overt responses would ultimately be controlled by the *crucial* element of each insigne. An *S* without insight would, however, be more likely to respond to *irrelevant* aspects of the stimulus-objects since the appropriate mediating responses would presumably be weaker.

In contrast with the present results, Lacey, Smith, and Green found that an informed-aware group (*S*s who had been informed that the word cow would be followed by shock, and were then subjected to the experimental procedure described above) showed *less* generalization from the word cow to other rural words than did the *S*s who were not informed. In this case, insight served to *reduce* mediated similarity. These data are, however, consistent with the present emphasis on the cue-producing function of insight. For the insightful *S*s, the word *cow* probably led to an implicitly verbalized expectation, such as: 'shock will follow.' After a series of trials, feedback from this response may have come to function as a part of the *CS* for the observed increase in heart rate. As a result, whenever a 'rural' word failed to elicit the insightful, 'shock-will-follow' response, part of the *CS* would be absent, and heart rate would remain relatively unaffected. For an *S* without insight, this response-produced cue would not be available as an aid in differentiating between the word *cow* and other 'rural' words. This interpretation is closely related to Dollard and Miller's assertion that "attaching distinctive cue-producing responses to similar stimulus-objects tends to increase their distinctiveness."[15]

SUMMARY

Twenty-three *S*s learned to discriminate between stimulus-objects representing 'friendly' and 'enemy' airplanes. These objects consisted of the Arabic numbers '2,' '3,' '4,' and '5,' arranged in a cross-shaped pattern. Although the same numbers appeared on each slide, their positions varied from one stimulus-object to the next, with the restriction that the number '2' (for friendly planes) or '5' (for enemy planes) always appeared on the left.

[15] John Dollard, and N. E. Miller, *Personality and Psychotherapy*, 1950, 101.

After this task had been mastered, generalization was tested by presenting *S*s with a new set of stimulus-objects, in which the Arabic numerals were replaced by different numbers of circles (*e.g.* two circles for the Arabic '2,' three circles for '3,' etc.); they were instructed to follow the same scheme that they had used in the initial series. The *S*s were then tested for retention with the original set of objects, and were subjected to a brief interview, to determine whether they had insight into the principle of the discrimination. A measure of each *S*'s intellectual ability was also obtained.

In the test for generalization, the *S*s without insight performed at a level which was significantly superior to the performance that would have been predicted from their verbalizations. The *S*s with insight were, however, more successful in responding to the test for generalization than were those without insight; this finding was not attributable to differences in *IQ,* learning speed, or retention, and seemed instead, to be a reflection of the cue-producing function of insight.

INFORMATIONAL REQUIREMENTS IN MAKING DECISIONS

By DEAN G. PRUITT, Yale University

One scheme which appears to be useful for analyzing decisions divides the process into four stages. When a person is faced with a problem, he will: (1) devise several alternative solutions, each of which seems to have some face validity; (2) then collect information on the good and bad aspects of each solution; (3) 'rate' each solution on the basis of this information; and (4) decide on the solution with the highest rating (the 'optimal' solution).[1] Most scientists who have adopted this model have been interested in the third stage, at which solutions are rated. Many feel that the 'expected-value' model of rational decision is a good description of what happens in this stage.[2] The present study is, however, focused on the second phase of the process—the collection of information relevant to the solutions being considered.[3] More specifically, the topic of this paper is the amount of information required before reaching a decision.

Fortunately for the psychologist, there has been, over a period of several hundred years, a strong interest among mathematicians, economists,

* Received for publication March 14, 1960. This report is derived from a thesis submitted in partial fulfillment of the requirements for the Ph.D. degree at Yale University.

[1] The value of this model has been challenged by March, Simon, and Guetzkow, who contend that it is applicable "only in exceptional cases." Instead of this 'optimizing' model, they propose a 'satisficing' model which states that people evaluate solutions serially and decide on the first solution encountered that has a satisfactory rating. See J. G. March, H. A. Simon, and Harold Guetzkow, *Organizations*, 1958, 140-141. In the view of the present author, both 'optimizing' and 'satisficing' are commonly found in human behavior. A pressing problem for research is to identify the circumstances under which each occurs. Since the 'optimizing' model is better developed than the 'satisficing' model, it was used in deriving the 'rational' strategies for this experiment.

[2] Ward Edwards, The prediction of decisions among bets, *J. exp. Psychol.*, 50, 1955, 201-214; C. H. Coombs and David Beardslee, On decision-making under uncertainty, in R. M. Thrall, C. H. Coombs, and R. L. Davis (eds.), *Decision Processes*, 1954, 255-286; Donald Davidson, Patrick Suppes and Sidney Siegel, *Decision Making: An Experimental Approach*, 1957, 1-2. Another approach involving an analysis of the concept 'risk,' has been suggested by C. H. Coombs and D. G. Pruitt, A study of decision making under risk, *J. exp. Psychol.*, 60, 1960, 265-277; Some characteristics of choice behavior in risky situations, *An. N.Y. Acad. Sci.*, 89, 1961, 784-794.

[3] For other treatments of this problem, see G. M. Becker, Sequential decision making: Wald's model and estimates of parameters, *J. exp. Psychol.*, 55, 1958, 628-636; F. W. Irwin and W. A. S. Smith, Further tests of theories of decision in an 'expanded judgment' situation, *ibid.*, 52, 1956, 345-348; Value, cost and information as determiners of decision, *ibid.*, 54, 1957, 229-232.

and statisticians in the rationale of making decisions. This work has provided a number of models which can be used, as first approximations to theory, for the development of hypotheses. As mentioned above, the 'expected-value' model of rational decisions has been very useful to researchers studying how people rate solutions. The same model was used in the present research: to predict how much information people will require before making a decision.

The experiment was designed to do two things: (1) to compare actual informational requirements with those predicted by the expected-value model; and (2) to compare the amount of information required before decision in two situations, *i.e.* (a) the 'predecisional,' in which no decision has been made and the individual is debating whether to choose solution *A* or *B*, and (b) the 'postdecisional,' in which the individual has already chosen *A* and is debating whether to reverse himself and choose *B*.

METHOD

Subjects. Twenty-four undergraduate men were used, all volunteers. Their ages ranged from 17 to 21 yr., with a median of 19 yr.

Procedure. *S* was seated before a board which had at the top red and green jewel-lights mounted side by side, and in the center a white warning light. *E* sat on the other side of the board. Every time the warning light flashed on, *S* could obtain an item of information (flash of the red or green light) by pushing a button. Counters mounted under the lights helped *S* keep track of how many of each kind of light he had seen. The *S* was told that a biased random process determined whether the red or green light would come on and was shown a device which purported to control this process. He understood that *E* could adjust this device to either of two settings, *A*, giving the red light a 60% chance of flashing and the green light a 40% chance, or *B*, giving the green light a 60% chance and the red light a 40% chance. In reality, *E* controlled the sequence of lights, which was divided roughly 70-30 or 30-70 between red and green (in some of the sequences red was predominant, in others green).[4]

S's task was to decide, for each sequence of lights, which setting the random device was in, whether biased toward red or green. There were two main conditions: (a) The Predecisional Condition, in which *S* could make his decision at any time during the sequence of items of information but was not permitted to change it; and (b) the Postdecisional Condition, in which *S* was required at the outset to decide in favor of one of the alternatives and could change this decision at any time during the sequence of items of information but could change it only once.

In the Predecisional Condition, *S* was given four problems of the following kind. First *E* announced that he had set the machine and that he was just as likely to

[4] The first Postdecisional Problem, which was for training purposes, utilized a sequence of 100 lights chosen from a table of random numbers with a 60-40 bias. The sequence was so rigged that *S*'s original choice was always verified by the trend of the lights.

have set it at *A* as at *B*. *S* pressed a single button once to obtain each item of information (red or green light). When he wished to decide on the solution to the problem (*A* or *B*) he stopped pressing the button and announced his decision. A bogus scoring system was introduced which provided incentives both for making a rapid decision and for making the correct one: The score for a correct decision was 101 points minus the number of items of information required before decision; no points were to be given for an incorrect decision. At the beginning of the experiment, *S* was told about the scoring system, "In this experiment I am studying how good people are at making decisions," and after the scoring system had been presented, he was informed that, "At the end of the hour, I'll tell you how well you did and how it compares with the average for the students at [this] university." Scores were not revealed until the end of the experiment.

In the postdecisional Condition, *S* was again given four problems. At the beginning of each problem he was required to guess whether the device was in Setting *A* or Setting *B*. Two buttons were mounted on the base of the panel. If he guessed *A*, he pushed the left-hand one to obtain information; if he guessed *B*, he pushed the right-hand one. He was permitted to change this decision at any time by shifting buttons but could change it only once. If he did not change his decision, he continued to receive information for a full 100 trials. If he did change it, the series was terminated, to prevent his learning about the balance of the series, knowledge of which might have served as a positive or negative reinforcement of his decision and thus introduced an unwanted factor of learning. A bogus scoring system was also introduced into this condition: Every time *S* pushed the correct button, he scored 100 minus *X* points, where *X* was the number of previous items of information received. His score was the sum of these points.

The crucial datum in the Predecisional Condition was the amount of information received before decision. In the Postdecisional Condition, the information in three of the four problems was so manipulated that *S* appeared to have made the wrong initial decision. The crucial datum in these problems was the amount of information received before changing decision.

Design. The same problems (*i.e.* sequences of informational lights) were used for all *S*s. Of the four Predecisional Problems given every *S*, two had sequences predominating in red lights and two in green lights. In the second, third, and fourth Postdecisional Problems, the sequences of information were so rigged that they increasingly counterindicated *S*'s initial decision. Two of the Predecisional Problems were exactly the same as two of the Postdecisional problems hence comparisons could be made in the amount of information required for decision with informational input held constant. Within the conditions, the order of problems and the pairing of problem with predominant color were completely counter-balanced. The Postdecisional Condition was always presented before the Predecisional Condition, the advantage of this design being that it gave *S* a chance to observe a long series of lights on the first problem, which was always a Postdecisional Problem in which the initial decision did not have to be changed. This procedure permitted *S* to become accustomed to the situation and to get a concrete notion of the nature of a random sequence with a 60-40 bias before the data were collected. As will be shown in the discussion of trends over time, the order of the two conditions cannot account for the differences which were found in informational requirements.

Predictions from the Model of Rational Decision

The expected-value model of decision was used to formulate hypotheses about the gathering of information under the assumption that problem solvers are periodically faced with the decision whether (a) to stop now and decide on a solution to the problem, or (b) to go on and collect more information. Any one of a myriad of strategies might be used in deciding when to stop. Since the expected values of all possible strategies could not be computed, a class of strategies was chosen for evaluation which appeared to make maximal use of the available information and, therefore, seemed most likely to include the optimal strategy, viz. taking information until one reaches a certain probability (p) of being correct about whether the apparatus is in Setting A or B. It was found that, in the tasks of the present experiment, the probability of being correct is a monotonic increasing function of the absolute difference ($|r - g|$) between the number of red and green lights seen.[5] Hence, expected values were determined (a) in the Predecisional Condition for the set of strategies h which may be described as "decide when $|r - g|$ reaches a certain value h" and (b) in the Postdecisional Condition for the set of strategies h which may be described as "change decision if $|r - g|$ reaches a certain value h and the proportion of red to green lights (*i.e.* the sign of $r - g$) counterindicates the initial decision."

The expected value for each strategy h was computed by a method analogous to the following.[6] All possible sequences of lights are enumerated and probabilities of occurrence under Settings A and B assigned to each. Then, for every trial, the sequences which reach $|r - g| = h$ for the first time are examined and the value of a correct decision multiplied by the proportion of cases in which this decision would be correct (sum of the probabilities of sequences in which the correct decision would be made). These products are then summed over all the trials to obtain an over-all expected value for the strategy. The results of these computations

---

[5] The probability of being correct if one chooses A when one has seen a majority of red, and B when one has seen a majority of green lights, is given by the following formula: $P = 1.5^{|r-g|}/(1 + 1.5^{|r-g|})$. It should be noted that the probability of being correct is also a function of the number of lights seen, $r + g$, since the probability that the more likely light will have been seen on a majority of the presentations increases as the number of presentations increases. Since, however, there is a stronger correlation between P and $|r - g|$ than between P and $r + g$, the former makes better use of the available information than the latter. Therefore, the optimal strategy was assumed to take $(r - g)$ into account.

[6] For a detailed description of this procedure see D. G. Pruitt, An exploratory study of individual differences in sequential decision making, Doctoral dissertation, Yale University, 1957, 5-16.

are shown in Table I. The maximal expected value for both conditions is found at $h = 4$, which was, therefore, assumed to be the rational strategy under both conditions.

RESULTS

All *S*s eventually decided in favor of *A* when red lights predominated in the series, and *B* when green lights predominated.

As mentioned above, the objective probability (and hence the uncertainty from the point of view of *S*) of being correct in a decision is a function of the absolute difference between the number of red and green lights seen before decision. This difference was, therefore, used as a rough index of the amount of information received before decision.[7] Because of

TABLE I

EXPECTED VALUES OF THE STRATEGIES EXAMINED

Strategy $h=\|r-g\|$	Probability of being correct p	Expected value in points: Predecisional Condition	Postdecisional Condition
0	.50	50.5	2525
1	.60	60.0	3295
2	.69	67.3	3694
3	.77	71.6	3871
4	.84	73.2*	3915*
5	.88	72.3	3882
6	.92		3807

* Largest expected value.

the skewness of the distribution of these values, a square-root transformation was made for all of the parametric tests of significance.

The mean, untransformed amount of information received in the Predecisional Condition was 4.46, and the median 3.50. A sign-test showed no significant difference from the optimal strategy of 4: 13 *S*s averaged less than 4, 2 averaged 4, and 9 averaged more than 4. The mean information required before decision in the Postdecisional Condition was 6.76 and the median 5.67. These averages are considerably *above* the optimal strategy of 4: 18 *S*s averaged above 4 and only 6 below 4, making the sign-test significant beyond the 5% level.

The results of the two conditions were compared in an over-all analysis of variance, which employed only the data from the two problems (schedules of lights) that appeared under both conditions. The difference be-

[7] Other indices which might have been used include such things as the probability of being correct or an information-theory measure of uncertainty, but all of these scales are equally arbitrary for this purpose.

tween conditions was highly significant ($F = 51.04$, 1 *df.*, $P < 0.01$), and there was a significant interaction between conditions and *S*s ($F = 6.57$, 23 *df.*, $P < 0.01$). The question arises whether the difference between conditions can be attributed to the fact that the Postdecisional Condition always preceded the Predecisional Condition. If such an order effect had been present, as a result, say, of fatigue or adaptation, we would expect to find it also operating on the replications within conditions. An analysis of variance for the Postdecisional Condition alone revealed that the position of a problem in time was not a significant source of variance ($F < 1$). A comparable analysis for the Predecisional Condition revealed a significant order-effect ($F = 6.16$, 3 *df.*, $P < 0.01$), but it was in the wrong direction for this explanation; the mean amount of information received in the first problem being 4.04, second problem 4.46, third problem 4.54, and fourth problem 4.83.

The latter two analyses of variance also showed very great differences among *S*s in the amount of information required before decision. (In the Postdecisional Condition, $F = 13.96$, 23 *df.*, $P < 0.01$; in the Predecisional Condition, $F = 33.73$, 23 *df.*, $P < 0.01$.) These individual differences were compared across conditions by means of a product-moment correlation computed on the *S*s' means. The coefficient was 0.51 ($P < 0.01$), indicating individual consistency in informational requirements across conditions.

Discussion

Both conditions in this experiment were designed to resemble situations in real life. An example of a predecisional situation would be a man trying to decide which of two houses to buy. A postdecisional situation would occur when the man has one house and is trying to decide whether to sell it and buy another. There is one point at which the experimental Postdecisional Problems do not strictly resemble this latter example, the absence of information *before* the initial decision. Without such information, the initial decision was necessarily more of a guess than a considered position. The Postdecisional Problems were started this way because of a desire to equate the amount of information at the beginning of these problems with that at the beginning of the Predecisional Problems. In the definitive aspect of real postdecisional problems, viz. the flow of information which progressively counterindicates a prior position, the problems of the experiment were, however, strictly comparable.

Comments from two *S*s suggest a reason for the difference in informational requirements under the two conditions. Both said that they had

taken more information in the Postdecisional than in the Predecisional Condition because the former permitted them to take information and have the possibility of scoring at the same time, whereas in the latter they could not score until they had stopped taking information. As one *S* put it, "You have a chance of being right and still testing in those (Postdecisional) problems." Stated more generally, we might conclude that people will require more information before changing a decision than before making one because of an impatience in the latter case to begin working toward goals.

SUMMARY

Twenty-four *S*s were required to decide whether a red or green light had the greater chance of flashing, using as information a sequence of flashes of the lights. In the Predecisional Condition, *S* could make this decision only once per sequence of lights. In the Postdecisional Condition, he was required to make an initial decision before seeing any lights and could change this decision once. A rational strategy concerning how much information should be taken before decision was derived for each condition on the basis of the expected-value formulation. In both cases the rational strategy was "continue taking information until the difference between the number of red and green lights reaches 4." The following results were obtained:

(1) In the Predecisional Condition, the amount of information required before decision closely approximated the rational strategy, but in the Postdecisional Condition, the amount of information required before changing decision exceeded the rational strategy.

(2) Although the two conditions were equated in informational input and rational strategy, considerably more information was required before changing decision in the Postdecisional Condition than before making decision in the Predecisional Condition.

(3) Within both conditions, the informational requirements for each *S* were highly consistent over replications. There was a significant, positive correlation between informational requirements under the two conditions.

PHOTOGRAPHIC MEASUREMENT OF THE RETINAL IMAGE

By ERIC G. HEINEMANN, Vassar College

This paper deals with the question of whether changes in accommodation and convergence affect the size of a retinal image. Suppose that two objects are presented at different distances from the eye. Let the objects be of such size that they subtend the same angle at a point located 1.35 mm. behind the vertex of the cornea (the first principal point of the unaccommodated eye), and assume accurate accommodation and convergence to each of the distances. Are the retinal images of the two objects of the same size?

Computational procedures based on the schematic eyes of physiological optics yield an unambiguous answer: The two images are very nearly of the same size. The image of the nearer object is very slightly smaller, because the first principal point shifts backward slightly in the accommodated eye.

There are a number of facts of psychophysics that may lead one to entertain some doubts about this matter. To give a few examples: In a recent paper, Heinemann, Tulving, and Nachmias showed that the apparent size of an object varies with changes in convergence and accommodation.[1] Objects look somewhat smaller when the eyes are adjusted for near vision than when they are adjusted for far vision. This effect does *not* depend on a correct perception of the distance of the objects. Furthermore, several investigators have reported that the size of the negative after-image (determined by measuring how much of the surface of projection it occludes) does not follow Emmert's law precisely; when the image is projected on a surface near the eye, its measured angular size is slightly larger than when it is projected on a surface that is farther away.[2] Analogous results for a figural after-effect have been reported by Sutherland.[3] None of these results is explicable in terms of current theories of size-constancy. In each case, the results are such as would be expected if accommodation or convergence produced a considerably larger decrease in the size

* Received for publication February 3, 1960. This work was done at the Psychological Laboratories, Harvard University, and was supported by a grant from the National Science Foundation.

[1] E. G. Heinemann, Endel Tulving, and Jacob Nachmias, The effect of oculomotor adjustments on apparent size, this JOURNAL, 72, 1959, 32-45.

[2] Wilhelm Schmülling, Aufdeckung latenter eidetischer Phänomene und des integrierten Typus mit der Intermittenzmethode, *Z. Psychol.*, 104, 1927, 233-321; Heinrich Klüver, reported in H. A. Carr, *An Introduction to Space Perception*, 1935, 366.

[3] N. S. Sutherland, Figural after-effect, retinal size, and apparent size, *Quart. J. exp. Psychol.*, 6, 1954, 35-44.

of the retinal image than that expected on the basis of the shifts in the cardinal points.

Is it altogether impossible that this is indeed what happens? The answer depends on how accurately the 'model' eyes of physiological optics represent actual human eyes. The optical constants of these 'model' eyes, *i.e.* the radii of curvature of the various refracting surfaces, their indices of refraction, and distances apart, represent average values based on empirical measurement of such constants in groups of living, normal, human eyes. The accuracy of the available measurements is high enough to raise no questions in the present context. Another quantity that must be known to compute the size of the retinal image is the length of the eyeball. The older measurements of the length of the eyeball all were based on excised eyes, but recently such measurements have been made on living eyes by a method that involves the use of X-rays.[4] The latter measurements were made, however, on eyes with accommodation relaxed. If the length of the eyeball were to change a little during the processes of accommodation and convergence, a change in the size of the retinal image would, of course, result. A survey of the relevant literature provides no evidence that would completely rule out this possibility. There are still other conceivable sources of variation in the size of the retinal image that cannot be dismissed on the basis of available evidence.

Whatever the plausibility of the particular speculations put forward on the basis of psychophysical results, some direct measurements bearing on the general question of whether the retinal image changes with accommodation and convergence are of obvious value. What follows is a description of a method for making such measurements and the results for a single *S*.

Subject. The *S* was a young man, 18 yr. old. His visual acuity, as determined on the Bausch and Lomb Orthorator, was equivalent to Snellen 20/20 for both near and far distances. The tests for phoria incorporated in the Orthorator showed no abnormalities, and retinoscopic measurements showed that, over the range of target-distances used in the experiment, changes in target-distance resulted in appropriate changes in the level of accommodation. (Retinoscopic measurements were made on the right eye only. The general technique used has been described elsewhere.)[5]

Apparatus. A diagram of the optical system used is shown in Fig. 1. The source of light is a projection-lamp with a small concentrated coiled filament (GE 18 AT 10). Light from this source is collimated by lens L_1. Lens L_2 forms an image of the filament in the plane of the pupil of *S*'s right eye. *X* is the slide that contains the pattern to be viewed by *S*. This slide is mounted about 1 mm. in front of Lens L_2. The pattern presented to *S*, at the viewing distance of 100 cm. (see below), is shown in Fig. 2, which is a photograph of its retinal image. The diameter of the whole illuminated area subtends 4° at a point 1.35 mm. behind the vertex of the cornea.

[4] S. Stenström, *Untersuchungen über die Variation und Kovariation der optischen Elemente des menschlichen Auges*, 1946; J. F. P. Deller, A. D. O'Connor, and A. Sorsby, X-ray measurement of the diameters of the living eye, *Proc. Roy. Soc. B*, 134, 1947, 456-467.

[5] Heinemann, Tulving, and Nachmias, *op. cit.*, 32-45.

The diameter of the largest of the three inner circles subtends 3°, the diameter of the next largest, 2°, and the diameter of the smallest, 1°. The black dot in the center is the point of fixation. The pattern presented at the viewing distance of 25 cm. was the same except that the three inner circles were omitted.

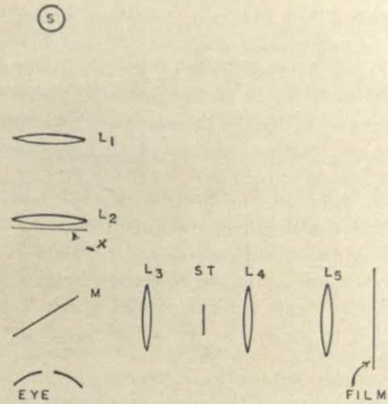


FIG. 1. DIAGRAM OF THE OPTICAL SYSTEM USED TO PHOTOGRAPH THE RETINAL IMAGE

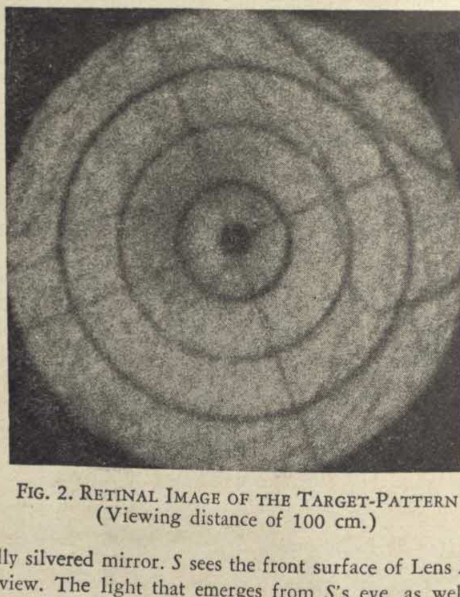


FIG. 2. RETINAL IMAGE OF THE TARGET-PATTERN
(Viewing distance of 100 cm.)

M is a partially silvered mirror. *S* sees the front surface of Lens L_2 (and the slide) in Maxwellian view. The light that emerges from *S*'s eye, as well as that reflected by the cornea, is reflected by the Mirror *M* toward Lens L_3. Moving in the direction of L_4, the Lens L_3 first forms an image of the retina, then an image of the corneal

reflection. A small stop (*ST*) is so placed in the plane where the image of the corneal reflection is formed that this light cannot reach the camera. Lens L_4 collimates the light from the retina. Finally, L_5, the objective of the camera, forms an image of the retina on the film.

The light-source and lenses L_1 and L_4 were so mounted on an optical bench that the distances could be adjusted easily. The position of *S*'s head was fixed by a biting board.

Procedure. *S*'s head was so adjusted that when he fixated the center of Lens L_4 with his right eye, this eye was in approximately its primary position. He always fixated binocularly, but, of course, only the right eye was presented with a Maxwellian view of the object.

For one series of photographs, the distance between the first principal point and the slide was adjusted to 25 cm., for another series this distance was adjusted to 100 cm. (To maintain constant angular size, the slides used at these two distances, of course, had to be of different linear size.) To facilitate precise adjustment of these distances, two sections of steel drill-rod were cut to lengths corresponding to the desired distances from the anterior surface of the cornea to the slide. For the final adjustment, one end of the rod was pressed against the slide, while the other end was pressed lightly against the cornea of the right eye.

S held in his hand a cable-release for the shutter of the camera. He was told to trip the shutter when he felt that he was fixating well and when the object appeared to be sharply focused. All exposures were 1/5-sec. in duration.

Altogether, about 200 photographs of the retina were made while *S* was viewing a target of the type described above. One-half of these were made with the target 25 cm. from the eye, while the other half were made with the target at a distance of 100 cm.

The main purpose of using a pattern consisting of concentric rings 1, 2, 3, and 4° in diameter was to provide a check on any distortions that might be introduced somewhere along the optical path from the stimulus-slide to the film. To provide somewhat more detailed information of the same sort, an additional series of 'calibration' photographs was made for a range of values around 4°. The targets used for this purpose were illuminated disks with a fixation-point in the center. The pattern is the same as that shown in Fig. 2 except that the three rings in the interior were omitted. Eight such targets were used, forming a graded series varying in diameter in successive steps of 5% from 3° 17′ to 4° 48′. The viewing distance in this part of the experiment was 100 cm.

Accuracy of measurement. The stimulus-disks having an angular size of 4° and also the additional disks used in the calibration-series were made by cutting holes in aluminum plates. These holes were cut to a tolerance of 1/200-in. in the value of the diameter. At the distance of 25 cm., this means a tolerance somewhat smaller than 1% for the disk of 4°.

Specification of the accuracy of the distance-measurements presents some special problems. The drill-rod was cut to the desired length with a tolerance of a small fraction of 1% but, even with the use of the biting board, *S*'s head is not, of course, completely immobile. The range of variation that might result from head-movements was determined roughly by observing the vertex of the cornea with an arrangement of a telescope and a measuring reticle as *S* tried to move his head up or down, or forward or backward, while maintaining a firm bite on the biting

board. The measured displacements that resulted did not exceed 1 mm. A quite conservative estimate is that, for the critical distance of 25 cm., the actual distance was within 1% (2.5 mm.) of the specified distance.

Errors in the diameter of the retinal image arising from uncertainties about the size of the stimulus-object and the viewing distance then should not exceed 2% of the specified value.

Treatment of results. Fig. 2 is a sample of the photographs obtained when the viewing distance was 100 cm. The picture shows the stimulus-pattern together with various structural features of the retina. The clarity of various features of the network of blood-vessels varied somewhat from photograph to photograph. The approximately 100 photographs taken at each distance first were carefully inspected. *E* selected

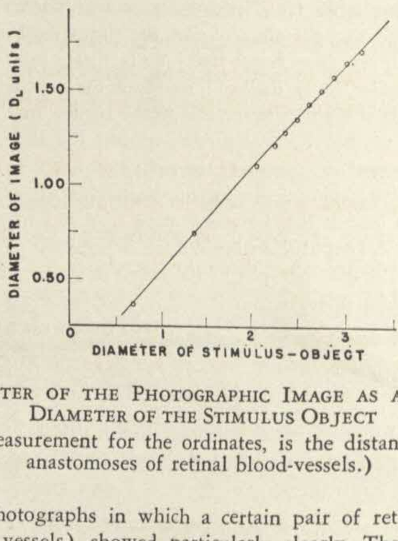


FIG. 3. THE DIAMETER OF THE PHOTOGRAPHIC IMAGE AS A FUNCTION OF THE DIAMETER OF THE STIMULUS OBJECT

(D_L, the unit of measurement for the ordinates, is the distance between a pair of anastomoses of retinal blood-vessels.)

from each set 18 photographs in which a certain pair of retinal 'landmarks' (anastomoses of blood vessels) showed particularly clearly. The two 'landmarks' selected were separated by an angular distance of about 3° and were so situated that a line connecting them passed near the point of fixation.

Measurements of size were made by the following procedure: Using a pair of dividers, *E* measured on each of the photographs the distance between the two 'landmarks.' Next, *E* drew on each photograph a straight line that was parallel to the line connecting the two 'landmarks' and that passed through the center of the point of fixation. Then *E* measured the distance between the two points where this line crossed the edges of the 4° illuminated area, and also, for the photographs taken at 100 cm., the distances between the points where this line crossed the three inner circles. Each of these distances then was divided by the distance between the 'landmarks.' This distance between the 'landmarks' will be called D_L. D_L then is the unit in terms of which the measurements of the size of the retinal image finally are expressed. Analogous measurements were made on the calibration-photographs.

Results. The measurements made on the calibration-photographs are shown in Fig. 3. The diameter of the stimulus-objects is plotted against the diameter of the photographs of their retinal images. The linear relationship obtained indicates that no significant distortions occurred.

Turning to the principal series of measurements, the mean diameter of the 4° area was 1.489, in D_L units, for the 100 cm. condition, with $SE_M = 0.0021$. It was 1.513 D_L units for the 25 cm. condition, with $SE_M = 0.0039$. The difference between these two means is significant beyond the 1% level of confidence as evaluated by the *t*-test.

The diameter of the image of the far object is 98.4% of the diameter of the image of the near object. While, as the *t*-test shows, the difference in the size of the images on the photographs is almost certainly real, the difference of 1.6% does not exceed the error that could occur as a result of inaccuracies in the measurement of the size of the stimulus-object and its distance.

The results then are in complete accordance with expectations derived from the 'schematic' eyes upon which computations of the size of the retinal image commonly are based. A solution of the problems raised by the psychophysical results mentioned in the introductory section is not to be found at the level of optics.

Summary. The image of a target-pattern of constant angular size formed on the retina of a living, human eye was photographed under two levels of accommodation and convergence. The size of the image was the same, within the limits of an error of measurement of about 2%, when the eyes were adjusted to a viewing distance of 25 cm. and to a viewing distance of 100 cm.

The results of these empirical measurements are in complete accordance with expectations based on the 'model' eyes commonly used in computations of retinal-image size.

LEARNING AND RECALL IN BILINGUALS

By SUSAN M. ERVIN, University of California, Berkeley

Natural bilinguals supply an opportunity to explore the effects of a type of verbal training too prolonged and complex to be easily duplicated in a laboratory. In the following study, the effects on memory of varying the languages used in learning and recall are investigated.

A bilingual may be said to be dominant in the language in which he has greater facility in naming common objects. One might expect that recall would be most complete in the dominant language. It is not obvious, however, that this superiority would be true for material learned in another language and recalled later in the language then dominant.

A model for bilingual recall is suggested by studies of mediated associations. Bastian has shown that the learning of *eagle–sickness,* for instance, facilitates later acquisition of *eagle–health,* presumably through the covert mediation of *sickness–health.*[1] Similar response-mediation can occur during naming of non-verbal stimuli. In a study of color-naming, the author showed that the names offered by bilinguals differ from those given by monolinguals, and that these were the names with the shortest response-time, or their translations.[2] A covert response seems to mediate a translation when response-language is restricted by instructions.

Similarly, it may be expected that in the naming of pictured items it will be found that: (1) the language of shorter reaction-time is the language of covert response when overt language is restricted; (2) covert responses in a different language from that used in the overt responses reduce recall in the latter language; (3) spontaneous translation is more probable into the dominant language than into the subordinate language.

In the following discussion, major items are those normally named more quickly in the dominant language and minor items are those named more quickly in the subordinate language.

The first two hypotheses lead to the prediction that major items will be recalled best when named in the dominant language and recalled in that language, as in Sector D of Table I. Similarly, for minor items recall will be facilitated in Sector

* Received for publication March 17, 1960. This research was supported by the Social Science Research Council as part of the 1955-56 Southwestern Project in Comparative Psycholinguistics, directed by John B. Carroll.

[1] J. R. Bastian, Response chaining in verbal transfer, Minnesota studies in the role of language in behavior, Technical Report No. 13, 1957, 44-45.

[2] Susan Ervin, Semantic shift in bilingualism, this JOURNAL, 74, 1961, 233-241.

TABLE I

PREDICTED RECALL UNDER DIFFERENT CONDITIONS

('Minor' and 'Major' refer to the type of items learned. Conditions marked 'plus' are above the mean of all conditions.)

Language of recall	Naming language: subordinate	Naming language: dominant
Subordinate	A Minor+	B Minor− Major−
Dominant	C Minor+ Major+	D Major+

A. If covert practice had no effect, then there would be no difference between major and minor items.

Spontaneous translation into the dominant language at the time of recall would facilitate recall in Sector C for both kinds of items. When would recall be worst? Presumably in Sector B, where items must be learned in one language and recalled in another, without the help of translation into the dominant language. This condition should be especially poor for major items, which are not even covertly practiced in the language of recall in Sector B. When items are pooled, Sector C should be optimal and B worst.

The above predictions might be compared to those based on other premises. If skill in the response-language alone mattered, then dominant recall should always be superior. If overt practice alone mattered, then the conditions in which the language is switched—Sectors B and C—would always be inferior. In neither case would major and minor items be different.

In a preliminary experiment on 16 'English-dominant' and 16 'Navaho-dominant' Navahos, tested on the reservation, sets of pictures were named and recalled under the four conditions described in Table I. The pictures were half of Indian and half of Anglo objects. It was found that better recall occurred whenever *S* was asked to recall in his dominant language ($p < 0.05$). Sector B was worse than all other conditions combined ($p < 0.03$) and Sector C was better than the other conditions combined ($p < 0.05$). Since *Ss* paced themselves, they were exposed longer to the items most difficult to name in the designated language. This varying exposure-time makes interpretation somewhat difficult. It was, however, clear that the pattern of ease of recall for the Navaho-dominant was the mirror-image of that of the English-dominant. Thus the pattern was the same regardless of which particular language was dominant.

METHOD

General. In the experiment reported below Italian-Americans were first given tests to establish control over individual differences. After an initial interview, they were given a language-dominance test to determine their relative skill in Italian and English, and an intrusion-test of spontaneous translation in recall. The learning experiment took place several weeks after the initial tests. The *Ss*, 32 in number, were assigned in a factorial design, according to dominant language, high or low intrusion-score, language of learning in the experimental session, and lan-

guage of recall in the experimental session. There were two *S*s in each cell of this design.

Language-dominance test. A set of 120 drawings of simple objects namable in both Italian and English was presented in a six-page booklet to each *S*. He was instructed to name the first three pages in English, then the next three in Italian. Then the procedure was repeated, reversing the order. The total response-time in Italian and in English for each set of three pages was recorded, together with the number of items *S* could not name. The difference in the response-times in log seconds for the naming of the same sets of pictures in English and in Italian constituted the test of relative dominance. As a validation, a similar test constructed for Navahos had yielded a correlation of 0.72 between English dominance and years of schooling. Schooling is the principal, but not the only source of English training for Navahos. In addition, other evidence on the background of the Italians supported the validity of the test.

Intrusion-test. To permit control over individual differences in the probability of spontaneous translation between the two languages, a further test was constructed. Six pages of five pictures each were presented, and *S*s named alternate pages in each language. At the end, they were asked to recall as many pictures as possible. The proportion of switches to the other language than the one in which the picture was named was used as a measure of intrusions for each language. This measure was used for a balanced design, a high-intrusion and a low-intrusion *S* being assigned to each experimental condition.

Learning. To guarantee that all *S*s could name the pictures readily in both languages, the 18 pictures in the learning experiment were chosen from the 120-picture test of language-dominance. Only those pictures were selected which all *S*s had named on both languages. This criterion guaranteed that the pictures were unambiguous, and that the terms were all common in both languages. While the actual exposure-time during naming in the test of language-dominance was not controlled, it is highly probable that these items, being the easiest, were named the most rapidly of the items in the test, hence effects on the learning task several weeks later would be minimal.

At the learning session, *S* was instructed to name the pictures in a specified language, and told that he would be asked to recall them later. Instructions were in English, due to the dialectal variations of Italian, but *S*s had heard *E* speak standard Italian. The series was in the same order for all *S*s, and was shown twice, with a 6-sec. interval between items. The response-time for naming each item was recorded. During the interval between learning and recall, Italian adjectives were assigned to Italian nonsense-words. After 6-min., *S*s were instructed to recall as many items as they could in a specified language. There was no time-limit on the period of recall.

Classification of items. After the collection of the data, the items which were easier to name in Italian than in English were identified empirically. For the Italian-dominant *S*s, six items were found which had shorter latencies in Italian than in English. These were Items 2, 5, 7, 11, 14, and 15 of the 18-item list, The six items named relatively faster in English were Items 3, 4, 6, 10, 12, 13. This categorization was based on a comparison of average latencies for the same items for the group which named in Italian and the group which named in English. Within each group, when the rank-order of latencies for different items was compared, the order was appropriate—for example, the Italian items had the shortest

latencies for the group speaking Italian and the English items the longest. There was a rank-order correlation of 0.60 between the latencies of these 12 items in the group speaking English and in that speaking Italian. This consistency suggested that a stable set of major and of minor items had been identified. These six items in each group were used in the analysis of results reported below.

In the English-dominant group, only three items were named more quickly in Italian than in English, and the rank-order of latencies of the items within the two groups was uncorrelated. No separate analysis of major and minor items was, therefore, possible for the English-dominant *Ss*.

The design did not counterbalance order of pictures. Since the major pictures were farther from the middle of the list for the Italian-dominants than were the minor pictures, they should have been remembered better, if the items were otherwise equated.[3] The predictions at issue concerned comparisons between conditions within each set of items separately, however, and across conditions serial order of items was constant.

Subjects. Interviews and tests were given to Italian-Americans in Boston, aged 20–80 yr. The extreme groups in dominance in the two languages were selected to comprise two groups of 16 *Ss*.

It was found that the two groups differed in many important respects. The Italian-dominants were older. None had come to this country before the age of 9 yr. They spoke Italian most of the time. None of the English-dominants came to this country after the age of 6 yr. and most were born here. They spoke English in school, with friends and siblings, learning Italian from older relatives but mixing it liberally with English. They were thus compound bilinguals whereas the Italian-dominants, who learned Italian from monolinguals, were more coördinate and thus less likely to have associative links between the languages.[4]

RESULTS

There were no significant differences in recall under the various conditions for the English-dominant *Ss*. The Italian-dominant *Ss* did however, display differential recall under various conditions. Minor (English) items were most often recalled when they had been named in English, regardless of the language of recall, ($F = 8.8$). Major items were most often recalled in the dominant language, Italian ($F = 9.1$), and there was an interaction between the language of learning and of recall ($F = 5.8$), with language-switching reducing recall.[5] These differences are significant at the 5% level.

[3] James Deese, Serial organization in the recall of disconnected items, *Psychol. Reports*, 3, 1957, 577-582.

[4] Susan Ervin and C. E. Osgood, Second language learning and bilingualism, in C. E. Osgood and Thomas Sebeok (eds.) Psycholinguistics, *J. abnorm. soc. Psychol., Suppl.*, 49, 1954, 139-146; W. E. Lambert, J. Havelka, and C. Crosby, The influence of language-acquisition contexts on bilingualism, *J. abnorm. soc. Psychol.*, 56, 1958, 239-244.

[5] In the analysis of variance, third-order interactions and within-*S* variance were pooled as an estimate of residual variance. For the resulting 1/9 *df*., $F_{0.05} = 5.12$, and $F_{0.01} = 10.56$.

Discussion

The results have supported the analysis of covert naming during learning. If covert naming were unimportant, there would be similar patterns of facilitation for major and minor items. In the Italian-dominant group, however, Condition A was facilitative only for minor items and Condition D for major items (see Table I). If there is only an overt effect of practice there should have been equal facilitation for all items in both conditions.

TABLE II

Mean Number of Items Recalled

Recall	English-dominant ($N=16$)		Italian-dominant ($N=16$)	
	Sub. naming	Dom. naming	Sub. naming	Dom. naming
Subordinate	9.50	11.25	7.50	5.50
Dominant	9.00	9.25	9.00	6.75

The significantly lower recall of major items in Condition B also demonstrates the presence of covert naming, since this is the only situation in which there is assumed to be neither covert practice in the language of recall nor translation.

TABLE III

Mean Number of Major and Minor Items Recalled

(Ambiguous items were excluded from this analysis, hence the totals do not equal those in Table II.)

(Italian-dominant *S*s)

Recall	Minor items		Major items	
	Sub. naming	Dom. naming	Sub. naming	Dom. naming
Subordinate	2.50	1.50	2.25	1.25
Dominant	3.75	1.75	2.50	3.50

The differences were significant only for the Italian-dominant bilinguals, and in the pilot study, for the Navahos, who had learned and used the languages under 'coördinate' conditions, separating the audiences and settings of use. Such conditions maximize of differences meaning,[6] and minimize the probability of intrusions.[7] For the English-dominant compound bilinguals, however, strong associative relationships between the two languages were likely as a result of the fact that the languages were used

[6] Lambert, Havelka, and Crosby, *op. cit.*, 242.

[7] Discrepancy of skill in two languages, as assessed by a test of language-dominance, is only reflected in intrusions in speech where there is a relatively frequent conversation with bilinguals, according to evidence of Ervin, The verbal behavior of bilinguals: The effect of language of report upon the Thematic Apperception-Test stories of adult French bilinguals, Unpublished Doctoral dissertation, University of Michigan, 1955, 155.

interchangeably with the same speakers and situations. Practice in one language then induces practice in the other by mediated generalization, and consequently recall-conditions would not differ.

This finding places an important restriction on the generality of language-effects on recall. Language will differentially influence recall only if the speaker addresses monolinguals frequently enough so that there is minimal probability of cross-language association.

Would these findings also apply to incidental learning? Probably so, since there is greater generalization under intentional learning.[8] Response-generalization extending to the other language would destroy the differential effects found in this study, so it seems reasonable to assume that under incidental conditions the differences would be at least as great. Indeed, it has been reported by Doob that for verbal statements there is differential recall under conditions of incidental learning, when languages of learning and recall are varied.[9]

SUMMARY

Italian bilinguals were tested for recall of pictorial material using English and Italian during learning and during recall. Pictures easier to name in the bilingual's more fluent language were recalled significantly more often in that language regardless of the language of learning. Optimal circumstances for recall of such pictures were learning and recall in the dominant language, if exposure-time during learning was controlled. The worst condition for recall was learning in the dominant language and recall in the other language.

Pictures easier to name in the bilingual's less fluent language were recalled equally well in either language, but learning in the subordinate language was superior.

Thus it appears that under natural conditions, where material would be learned in the lanuguage most appropriate to it, the optimal recall-language is always the language dominant at the time of recall. These differences between conditions appeared only for coördinate bilinguals, who had had an opportunity to learn both languages from monolinguals.

These differences fit the following assumptions: (a) there are covert responses in the language of easier naming, when overt language is restricted; (b) covert responses in the same language as the overt response strengthen recall in that language; (c) spontaneous translation is more probable into the dominant language.

---

[8] Leo Postman and P. A. Adams, Studies in incidental learning: VI. Intraserial interference, *J. exp. Psychol.*, 54, 1957, 153-167.

[9] L. W. Doob, The effect of language on verbal expression and recall, *Amer. Anthropol.*, 59, 1957, 88-100.

THE EFFECT OF ISOLATION OF STIMULI AND RESPONSES IN PAIRED ASSOCIATES

By Jacob Nachmias, Henry Gleitman, and Virgil V. McKenna, Swarthmore College

Beginning with Von Restorff, numerous investigators have shown that a distinctive item is learned more readily than relatively undifferentiated items in the same list.[1] Although this 'isolation-effect' has been treated variously in terms of Gestalt trace-theory, of surprise, or of reduced intraserial interference, no generally acceptable explanation of the phenomenon exists at present.[2] A necessary prerequisite for such a theory is a substantial body of data regarding the properties of this effect. This prerequisite has not yet been met, despite a recent recurrence of interest in the phenomenon.

Thus, it is important to know the relative efficacy of isolation of the stimulus and the response in associative learning. Among other things, such information would be relevant to the evaluation of theories that explain the isolation-effect as primarily due to greater stimulus-generalization among the crowded pairs. If so, the effect should be less on the side of the response.

A recent experiment by Kimble and Dufort bears on this problem.[3] They conclude that responses to isolated stimuli are learned more easily than control items, but isolated responses are learned, if anything, less easily. Unfortunately, this conclusion is not quite warranted, in view of a shortcoming in their experimental design. In their Experiment II—the most relevant one for this problem—10 common three-letter words were paired with 8 familiar two-syllable words and 2 paralogs (the isolated items) selected from Noble's list. For Group S, the three-letter words were the response (isolation occurring on the side of the stimulus), while for Group R they were stimuli (isolation occurring on the side of the response). Consequently, Groups S and R differ in two respects: isolation on either the side of the stimulus or of the response, and also the kind of material that served as stimuli and responses. Since it is quite likely that the magnitude of the isolation-effect depends

* Received for publication February 13, 1960. The authors are indebted to Charles C. Torrey for conducting some of the experiments in this study.

[1] H. von Restorff, Ueber die Wirkung von Berichtsbildung im Spurenfeld, *Psychol. Forsch.*, 18, 1933, 299-342.

[2] R. T. Green, Surprise as a factor in the Von Restorff effect, *J. exp. Psychol.*, 52, 1956, 340-344; E. J. Gibson, A systematic application of the concepts of generalization and differentiation to verbal learning, *Psychol. Rev.*, 47, 1940, 196-229.

[3] G. A. Kimble and R. H. Dufort, Meaningfulness and isolation as factors in verbal learning, *J. exp. Psychol.*, 50, 1955, 361-368.

on the type of response *S* must learn and on the type of stimulus to which he must associate them, Kimble and Dufort's design is not suitable for the comparison of S and R isolation. Experiments such as that of Newman and Saltz, using serial learning, are inconclusive for the same reason.[4]

The present investigation attempts to resolve the uncertainty regarding the relative efficacy of isolating S and R. To avoid confounding type of material with the variable under study, the identities of the isolated items, of the items in which they are embedded (crowded items), and of the items with which crowded and isolated items are paired (homogeneous items) are all systematically varied.

METHOD

Materials. Three categories of materials were employed: two-place numbers, nonsense syllables (from the 20% level of Glaze's list), and five-letter adjectives (rated 'A' in the Thorndike-Lorge word count).[5]

To balance the effect of materials for isolation of both stimulus and response, 12 different paired-associate lists were constructed. Every list consisted of nine pairs, presented in scrambled order. The stimulus and response of any pair never belonged to the same category of items. To study isolation of the stimulus, lists were used in which all responses were from one category, the stimuli being drawn from the other two—seven of them from one of these categories ('crowded items'), and two from the other ('isolated items'). To study isolation of the response, the same lists were used but with the pairs reversed, that is, stimulus and response being interchanged. To control for the effect of materials, isolated and crowded categories were reversed in different lists, so that the very same items that were 'isolated' in one list became 'crowded' in another. This procedure was followed for all three categories of items, counterbalancing materials for all conditions. The different lists are diagrammatically presented in Table I.

Procedure. Sixty *Ss* were employed in this experiment, five for each list. All *Ss* were undergraduates taking an introductory course in psychology, and were randomly assigned to one of the 12 groups.

Every list, in eight scrambled orders, was presented on a Hull-type memory-drum. The stimulus was presented for 1.5 sec., followed by both stimulus and response for another 1.5 sec. Three seconds elapsed between successive presentations of the list. The *Ss* were instructed to anticipate the response-item. Syllables had to be spelled and numbers were reported in digits. This procedure was continued to a criterion of one errorless run through the entire list.

RESULTS

Stimulus-isolation vs. response-isolation. The first analysis was based upon the comparison of pairs containing isolated items with pairs contain-

---

[4] S. E. Newman and Eli Saltz, Isolation effects: stimulus and response generalization as explanatory concepts, *J. exp. Psychol.*, 55, 1958, 467-472.

[5] J. A. Glaze, The association value of nonsense syllables, *J. genet. Psychol.*, 35, 1928, 253-267; E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

ing the so-called 'critical crowded' items, the latter being those crowded items in the list under consideration which were isolated in the other list. The data from the remaining five pairs were not used.

TABLE I
SCHEMATIC REPRESENTATION OF THE COMPOSITION OF THE TWELVE LISTS
(A=adjectives; N=numbers; S=nonsense-syllables)

Condition	Stimuli	Responses	Trials to criterion	
			Mean	*SD*
Stimulus-isolation	A, S	N	24.4	8.1
	A, N	S	37.8	18.8
	S, A	N	20.0	18.6
	S, N	A	18.6	2.9
	N, A	S	28.6	9.1
	N, S	A	16.2	8.8
Response-isolation	N	A, S	33.0	9.8
	S	A, N	32.0	27.7
	N	S, A	27.4	16.4
	A	S, N	20.0	9.8
	S	N, A	28.4	19.6
	A	N, S	26.8	10.5

Table II indicates the mean number of all errors (including failures to respond) for pairs containing isolated and 'critical crowded' items for both conditions of isolation. The data were subjected to an analysis of variance (Type I, mixed design) with two variables: items isolated vs. items crowded, and stimulus-items homogenous vs. response-items homogeneous.[6] A difference in the relative efficacy of the isolation of stimulus and response would show up as an interaction between these main variables. In fact, the analysis reveals a highly significant effect of isolation ($F = 26.0$, $df. = 1{,}58$; $p < 0.01$). Both the interaction, and the other main effect, fall, however, far short of statistical significance ($F < 1.00$).

To control for the possibility that these results are due to differences in difficulty of the lists as a whole, another analysis was performed, using the data from all items in the lists. For every *S*, the mean number of errors for both isolated and critical crowded items were corrected by subtracting the mean errors on the remaining five (non-critical crowded) pairs. An analysis of variance performed on these 'corrected' scores yielded results almost identical to those reported above: effect of isolation highly significant ($F = 29.2$, $df. = 1{,}58$; $p < 0.01$), the other main effect and interaction both insignificant.

It thus appears that isolation is effective *both* on the side of stimulus

[6] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 267-273.

and of the response of paired-associate lists, and to about the same extent.

Intrusion-errors. The preceding analysis was based upon all errors for pairs containing critical items, whether they were intrusions or failures to respond. One further fact emerges when intrusions are considered separately.

These errors account for a fairly large proportion of the total: 33% for pairs containing isolated and 27% for pairs containing 'critical crowded' items, respectively.

TABLE II

MEAN NUMBER OF ERRORS IN ISOLATED AND CRITICAL-CROWDED PAIRS

Condition	All errors: isolated pairs	All errors: critical-crowded pairs	Intrusion-errors: isolated pairs: within	Intrusion-errors: isolated pairs: across	Intrusion-errors: critical-crowded pairs: within	Intrusion-errors: critical-crowded pairs: across
Stimulus-isolation	21.3	28.8	3.5	5.3	1.2	8.4
Response-isolation	23.6	33.3	3.2	2.5	.8	5.6

When considering pairs with isolated items, two kinds of intrusions may be distinguished: *within-category,* and *across-category.* A *within-category* intrusion occurs when one isolated item is substituted for another (or the response to one isolated item is given in place of the response to another, depending on whether isolation is on the side of the stimulus or of the response). An *across-category* intrusion occurs when a crowded item (or the response to a crowded item) is substituted for an isolated item (or response to an isolated item).

Analogous measures can be devised for pairs containing 'critical crowded' items. A *within* intrusion becomes the substitution of one critical crowded item (or the response to it) for the other (or the response to it). The analogue to an *across* intrusion is giving an item other than a critical crowded one (or the response to one) for a critical crowded item (or the response to it).

As indicated in Table II, there are proportionately more *within* than *across* intrusions in pairs containing isolated items than in pairs containing critical crowded ones. When isolation was present on the side of the stimulus, the ratios of average *within* to average *across* intrusions are 0.65 and 0.15 for isolated and crowded pairs, respectively. Corresponding ratios for isolation of the responses are 1.28 and 0.14. The differences between such ratios for crowded and isolated pairs, based on data from individual *Ss,* were subjected to the sign-test, and found to be highly significant ($p <$ 0.01). Evidently the isolated items become more similar to each other by

virtue of their isolation, thus leading to proportionately greater confusion between them. It should be recalled that, considering all 12 lists, the same items appeared equally often in isolated and critical-crowded pairs.

As was the case when total errors were considered, no significant differences appear to exist between isolation of the stimulus and of the response in so far as the preponderant type of intrusion-error is concerned.

Summary

An experiment is reported which sought to determine the existence and relative magnitude of the effect of isolation on the side of the stimulus and of the response in paired-associate learning. The type of material from which were chosen the isolated items, the crowded items in which they were embedded, and the items with which isolated and crowded items were paired was systematically varied. It was found that isolated items (or responses to them) are learned more readily than crowded items (or responses to them), and that the efficacy of isolation is about the same on the side of the stimulus and of the response. Analysis of intrusion-errors revealed that there is relatively more confusion between isolated items (or responses to them) than between crowded items (or rsponses to them).

IMMEDIATE AND RESIDUAL FIGURAL AFTER-EFFECTS IN KINESTHESIS

By R. ALLEN GARDNER, Quartermaster Research and Engineering Center, Natick, Massachusetts

Most of the more careful studies of figural after-effects (*FAE*) have been concerned with transient changes, although the literature contains references to effects that persist for 25 hr. or more.[1] Recently, Wertheimer and Leventhal demonstrated a residual after-effect in kinesthetic judgments that appears to cumulate over at least four sessions of satiation spaced at intervals of 24 hr.[2] In view of the lack of previous work on this topic and the relevance of such residual effects to the effects of distributed practice in motor learning,[3] the present experiment was designed to replicate and expand upon the work of Wertheimer and Leventhal by a somewhat more refined method.

Apparatus. The standard and the variable were two pairs of metal disks, 2 cm. in diameter. They were so mounted that they represented the ends of two imaginary cylinders whose long axes lay on a common line which could be drawn parallel to the shoulders of *S* and 35 in. from the floor. The pairs of disks representing the standard and the variable were held simultaneously between the thumb and index fingers of the left and right hands, respectively. The two thumb-disks were fixed 30 cm. apart, while *E* could vary the length of either imaginary cylinder by moving the index-finger disks which were mounted on suitable slides. The index-finger disk of the variable also was coupled with an electric motor which *S* could operate by means of knee-switches. The switch operated by the left knee decreased and that operated by the right knee increased the length of the variable at the rate of 0.45 mm. per sec. A Vernier scale permitted *E* to read the settings to the nearest 0.1 mm. *S* sat facing the apparatus at a distance which he found comfortable, thrusting his hands through holes in a curtain which prevented visual inspection of the disks.

Procedure. Each *S* was tested five times at intervals of 24 hr. Each test consisted of 12 pre-satiational judgments and a single post-satiational judgment. The method of adjustment was used and the standard was set at 45 mm. for all judgments. For

* Received for publication March 3, 1960. The author is now at Wellesley College.

[1] C. P. Duncan, On the similarity between reactive inhibition and neural satiation, this JOURNAL, 69, 1956, 227-235; Peter McEwen, Figural after-effects, *Brit. J. Psychol., Monogr. Suppl.*, No. 31, 1958, 1-106.

[2] Michael Wertheimer and C. M. Leventhal, "Permanent" satiation phenomena with kinesthetic figural after-effects, *J. exp. Psychol.*, 55, 1958, 255-257.

[3] Duncan, *op. cit.*, 227-235; Duncan, Visual and kinesthetic components of reactive inhibition, this JOURNAL, 70, 1957, 616-619.

each of the pre-satiational judgments, *E* pre-set the variable by adding or subtracting 6 mm., 8 mm., or 10 mm. to or from the last setting made by *S* according to a predetermined sequence. The sequence of pre-settings was scrambled and changed each day with the restrictions that plus and minus pre-settings were in ABBA order and each of the six possible pre-settings was used once during the first six and once during the last six judgments. The first pre-setting on each day was plus for one half of the *S*s and minus for the other half. On Day 1, each *S* made four practice judgments and his setting for the last practice judgment was used to determine the pre-setting for his first experimental judgment. On each succeeding day, the pre-setting for the first judgment was determined from the last pre-satiational judgment made on the preceding day. The pre-setting for the single post-satiational judgment of each day was the median of the last six pre-satiational judgments made on that day.

Design. There was an interval of 4 min. between the last pre-satiational judgment and the single post-satiational judgment. During this time, the two experimental groups held the standard and the variable between their fingers as if making a judgment, but the electric motor was off and neither the standard nor the variable could be changed in length by *S*. For Group E(+) the standard and variable were set at 15 mm. and 75 mm., respectively. For Group E(−) these valves were reversed, while the *S*s of Group C merely sat in their places for 4 min. with their hands removed from the apparatus. Each of the three major groups was divided into two equal sub-groups; a Massed sub-group that made each of the 12 pre-satiational judgments consecutively without interruption, and a Spaced sub-group that had a 2-min. rest between the sixth and the seventh judgments.

Subjects. There were 60 *S*s, divided equally among the groups and sub-groups of the experiment. They were enlisted men stationed at the US Army Garrison, Fort Devens, Massachusetts.

Results. The Spaced and the Massed sub-groups were found to be indistinguishable from each other on all of the measures, including the difference between the first six and the second six pre-satiational judgements. The same was true of the sub-groups created by having half of the *S*s begin their daily sessions with an ascending judgment and half with a descending judgment. Consequently, only the three major groups, E (+), E (−) , and C, will be considered in the remaining discussion.

On each day it was possible to estimate the point of subjective equality (*PSE*) either from the first six judgments, the second six judgments, or all of them. The three measures were very nearly identical, and the levels of significance of any tests were not substantially altered by choosing one or the other of them. For precision, the following conventions were adopted for estimating the immediate and the residual *FAE*. For each day, immediate *FAE* was taken as the difference between the mean of the last six pre-satiational judgments and the single-post satiational judgment. The residual *FAE* was estimated by the *PSE* obtained from the means of the first six pre-satiational judgments of each day.

The daily immediate *FAE* is plotted in Fig. 1 for the three major groups. For Group C, the average value of this measure for all five days was + 0.23 mm., which is not significantly different from zero by the *t*-test. The differences between Group C and the two experimental groups on this measure were +6.18 mm. for Group E(+) and −5.94 for Group E(−). By Tukey's test for multiple comparisons, Group E(+) and Group E(−)

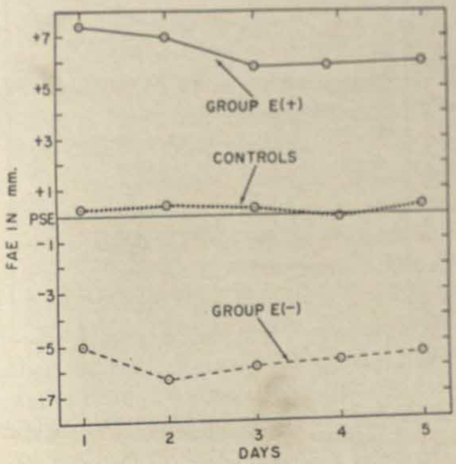


FIG. 1. IMMEDIATE *FAE*
(*PSE* is the mean of the last six pre-satiational judgments.)

were significantly different from each other and from Group C at the 5% level of confidence.[4] If the sign of the immediate *FAE* is reversed for Group E(−) to make comparison of absolute differences possible, the significance of the difference between the two experimental conditions vanishes. In Fig. 1 it can be seen that the greatest discrepancy between Group E(+) and Group E(−) in their absolute difference from Group C occurred on Day 1. Accordingly, a separate *t*-test was made for Day 1, but this, too, failed to show significance.

The daily residual *FAE* is plotted in Fig. 2 for the three groups. The slopes of these curves were estimated by computing regression-equations for the five days using the individual means of the first six pre-satiation judgments. The slopes were 1.77 for Group E(+), −0.35 for Group E(−), and 0.20 for Group C. The slope for Group E(+) was significantly different from zero with $P < 0.001$, but neither of the other two

[4] G. W. Snedecor, *Statistical Methods*, 1956, 251-253.

slopes was significant. The slope of Group E(+) was significantly different from that of Group C with $P < 0.003$, but the slope of Group E(−) was not reliably different from that of Group C.

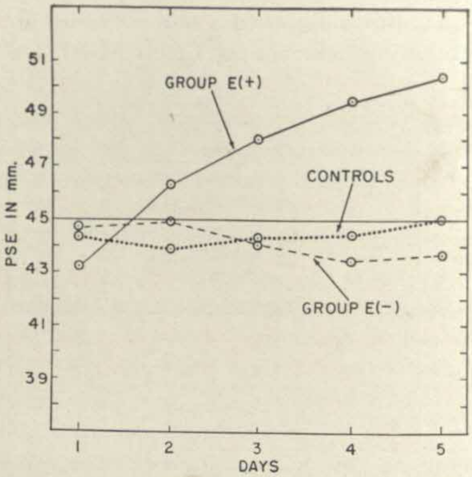


FIG. 2. RESIDUAL *FAE* AS SHOWN BY SHIFT IN *PSE*
(*PSE* is the mean of the first six pre-satiational judgments.)

With a procedure very similar to that for Group E(+), Wertheimer and Leventhal used 1½ in. (39 mm.) as a standard and satiated their *Ss* with lengths of ½ in. (13 mm.) in the left hand, and 2½ in. (65 mm.) in the right hand.[5] Wertheimer and Leventhal did not take a comparable measure of immediate *FAE*. For the residual *FAE*, as nearly as can be estimated from their Fig. 1, the increase in *PSE* that they obtained from Day 1 to Day 5 was between 7.0 and 7.2 mm. For Group E(+) in the present experiment, the cumulative effect on *PSE* from Day 1 to Day 5 was 7.2 mm. This would indicate that the various difference in procedure were negligible in their effects on the residual *FAE*.

Discussion. The residual *FAE* of Group E(+) represents a clear confirmation of the findings of Wertheimer and Leventhal for their similarly treated group. On the other hand, after equal but opposite satiation, Group E(−) showed equal but opposite immediate *FAE* with no detectable residual *FAE*. It is possible that the two conditions did differ in the amount of immediate *FAE* that was induced, but the measurable effects of satiation were at a maximum for both conditions when the immediate test was taken. A differential rate of decay such as that reported for visual *FAE* in

[5] Wertheimer and Leventhal, *op. cit.*, 255-257.

Sagara and Oyama,[6] might then be used to account for the asymmetrical *FAE* found in the residual tests. Köhler and Dinnerstein point out that in kinesthesis, as in vision, the Köhler-Wallach theory of *FAE* does not require that symmetrical effects be found after equal but opposite procedures of satiation.[7] If the theory demanded *either* symmetrical or asymmetrical results on *both* the immediate and the residual tests, it would be contradicted by the present results; but, if *both* symmetry and asymmetry on *either* test can be compatible with the theory, then it is by no means confirmed. Unless the differential conditions which yield now symmetrical and now asymmetrical *FAE* can be derived from the theory, its compatibility with the results both of Group E(+) and of Group E(−) merely serve to illustrate its invulnerability to empirical test.

The literature for visual *FAE* also furnishes conflicting findings on the persistence of *FAE*. Recently, Sagara and Oyama reviewed Japanese work indicating dissipation within 2 min., while Mountjoy reports detectable residual effects after at least 24 hr.[8] Numerous differences in technique would account for the discrepancies among the previous findings, but the present experiment seems to indicate that it may not be safe to generalize a finding regarding the presence or absence of residual *FAE* from one *I*-figure to its mirror image.

Summary. Each of 60 *S*s made 13 kinesthetic judgments on each of five days. Between the 12th and the 13th judgments, Group E(+) was satiated for 4 min. in such a way as to induce overestimation of a 45-mm. standard, Group E(−) was satiated to induce underestimation, and Group C merely rested for 4 min. On each day, immediate *FAE* was taken as the difference between the 13th judgment and the mean of judgments 7-12. The residual *FAE* was taken as the daily shift in *PSE* as measured by the mean of judgments 1–6. On the immediate test, the judgments of Group E(+) and those of Group E(−) were significantly different from each other and from the post-rest judgments of Group C, but the absolute magnitudes of overestimation and underestimation were very nearly the same. Group E(+) showed a significant cumulative residual effect on each day, while Group E(−) showed no detectable residual effect at any time during the experiment.

[6] Moriji Sagara and Tadasu Oyama, Experimental studies on figural after-effects in Japan, *Psychol. Bull.*, 54, 1957, 327-338.

[7] Wolfgang Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesis, *Miscell. Psychol. Albert Michotte,* 1947, 196-220; Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

[8] Sagara and Oyama, *op. cit.*, 327-338; P. T. Mountjoy, Effects of exposure-time and intertrial interval upon decrement to the Müller-Lyer illusion, *J. exp. Psychol.*, 56, 1958, 97-102.

THE EFFECT OF AN INCREASED VERSUS A DECREASED REDUCTION IN SHOCK USED AS INCENTIVE

By Edmund S. Howe, University of Maryland School of Medicine

According to Hull, the role of incentive (K) is that of increasing or decreasing the reaction-potential, ${}_{S}E_{R}$.[1] This implies that if the value of K be changed during performance of a well learned act, then ${}_{S}E_{R}$ will change rapidly to a magnitude typical for behavior that had been reinforced from the outset with this new value of K. This deduction from Hull's theory has been explored employing the hunger drive and changes in the magnitude of the food-incentive, but the deduction holds true only when the incentive is suddenly increased.[2] When the amount of incentive is suddenly *decreased* the ensuing value of ${}_{S}E_{R}$ has repeatedly been found to be *below* the theoretically expected value, a phenomenon known as the "Crespi Depression Effect," after the author who first systematically observed it.[3]

The experiment reported here was designed to study the prediction from Hull's model when electric shock is the motivating agent, and varied amounts of reduction in the shock constitute the incentive, K, and hence the reinforcement for performance of an instrumental response.

Method

Subjects. The *S*s were 13 Sprague-Dawley albino rats, approximately 110-130 days old at the outset of the experiment. They were on *ad libitum* food and water, but all food was removed from the cages some 5 hr. before, and was returned to the cage 1 hr. after, each day's trials.

Apparatus. The apparatus consisted of the electrified alley described in detail by Campbell and Kraeling.[4] Shock was given through a 250,000Ω resistance in series with *S*. The starting-box (12 in. long) and the center portion (36 in. long) were continuously electrified with 400 v. The goal-box (12 in. long) was continuously

* Received for publication March 26, 1959. The study was performed while the author was a Smith-Mundt post-doctoral fellow at Yale University.

[1] C. L. Hull, *A Behavior System*, 1952, 140-148.

[2] K. W. Spence, *Behavior Theory and Conditioning*, 1956, 127-148; R. S. Czeh, Response strength as a function of the magnitude of the incentive and consummatory time in the goal box, unpublished Master's thesis, University of Iowa, 1954. (Not seen; cited by Spence, *op. cit.*, 132).

[3] L. P. Crespi, Quantitative variation of incentive and performance in the white rat, this Journal, 55, 1942, 467-517; David Zeaman, Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 39, 1949, 446-483.

[4] B. A. Campbell and Doris Kraeling, Response strength as a function of drive level and amount of drive reduction, *J. exp. Psychol.*, 45, 1953, 97-101.

electrified with either 300 v. or 150 v., so that upon entering the goal-box S experienced a predetermined reduction in the strength of the shock. Running time was recorded on a Springfield timer which was set off by E's raising the door of the starting-box, and stopped by his pressing a toggle switch.

Procedure. The S was placed facing in the direction of the goal-box on every trial. The door of the starting-box was then raised, automatically setting off the timer. When S passed under a marker on the plexiglass roof of the goal-box, the timer was stopped, and the door of the goal-box dropped behind him. The S was then detained in the goal-box with some reduced drive-strength for 15 sec., before being removed to his cage. Each S was given five trials at approximately 15-min. intervals, on each of six consecutive days. For the first four days (the *pre-shift phase*) S was consistently given *either* a low reduction in shock (*i.e.* low incentive) *or* a high reduction in shock (*i.e.* high incentive) upon entering the goal-box. During the last two days' trials (the *post-shift phase*) the amount of reduction in shock in the goal-box was shifted either from the pre-shift low to a consistent post-shift high, or from pre-shift high to a consistent post-shift low.

Experimental design. Seven of the 13 Ss were in the low and 6 were in the high incentive group during the pre-shift phase (Trials 1-20). During the post-shift phase (Trials 21-30) the amount of incentive was reversed for the two groups. We may thus refer to a low-high and a high-low group.[5]

RESULTS

The mean running times, shown in Table I, are plotted in Fig. 1. Each value was obtained by finding first the median running time for each rat over a given five-trial block, and secondly, the mean of such medians for a given group of Ss over the same five-trial block. It will be observed that on the pre-shift trials the low-incentive group consistently ran more

TABLE I

MEAN RUNNING TIMES.

Each entry is the mean of individual medians over separate blocks of five trials.

Group	Blocks of Trials: pre-shift				post-shift	
	1–5	6–10	11–15	16–20	21–25	26–30
Low-high	4.85	1.64	1.74	1.47	1.45	1.41
SD	6.24	.18	.47	.24	.27	.18
High-low	1.59	1.31	1.15	1.12	1.44	1.35
SD	.14	.17	.15	.19	.38	.19

[5] The design does not include two control groups: one running the entire series with the high incentive, and the other running the entire series with the low incentive. Indeed, there were initially 16 Ss in a design including these control measures. Unfortunately, one S was injured in the alley, and two Ss developed unintended responses which enabled them to avoid shock in the starting-box. These Ss were therefore discarded. The data at the end of the pre-shift phase gave E an impression that the Ss were performing asymptotically, and a decision was made to pool the remaining Ss as described. The results give no reason to believe that E's impression was incorrect.

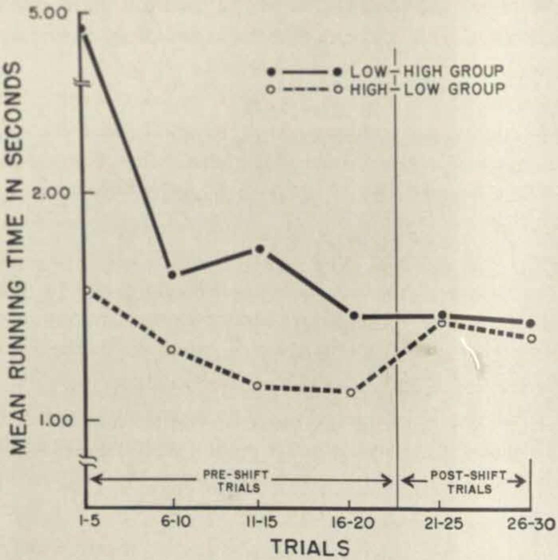


FIG. 1. MEAN RUNNING TIMES OBTAINED FROM INDIVIDUAL *Ss'* MEDIANS OVER FIVE-TRIAL BLOCKS.

slowly than the high-incentive group. An analysis of variance was performed to estimate the reliability of this difference. Since there were several extremly erratic running times during the first five-trial block, the two groups were compared only over the last fifteen pre-shift trials. With the mean square among *Ss* in the same group as the error term, the *F*-value between groups is 16.46 ($P < 0.01$). Furthermore, the significantly longer mean running time of the low-incentive group persisted during the last pre-shift block of five trials ($t = 3.27$). These results are in accordance with the prediction yielded by a drive-reduction theory: that performance level is some positive function of the amount of drive reduction.[6] They also accord with the earlier findings of Campbell and Kraeling who, with a broader range of drive-intensities, established that "running speeds vary systematically with the amount of drive reduction, throughout training as well as the limit of performance."[7]

[6] N. E. Miller, Learnable drives and rewards, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 435-472; Hull, *op. cit.*, 5-6; John Dollard and N. E. Miller, *Personality and Psychotherapy*, 1950, 39-42.

[7] Campbell and Kraeling, *op. cit.*, 99.

There were no significant changes in running time during the last 15 pre-shift trials. This suggests that the asymptotes of performance under the two incentive conditions had been reached.

Turn now to the post-shift performance levels. It is evident from Fig. 1 that the low-high group continued to perform at the same level even when the incentive was increased. No significant difference was found, so that it appears that introduction of a larger incentive does not rapidly evoke faster performance. The high-low group, on the other hand, shows a radical *increase* in running time after the introduction of a lower incentive. A comparison of running times of this group during the last block of pre-shift trials with the mean of individual median running times over the entire 10 post-shift trials, gave a t-value of 3.87. Finally, during the post-shift trials the high-low group increased its running time to approximately that of the low-high group. The comparison of the two groups over the 10 post-shift trials yielded a t-value of only 0.28, failing to show a real difference.

Discussion

The results of this experiment are in sharp contrast with the findings of those who have worked with the effects of sudden changes in the food-incentive. The effect of decreasing the incentive in this study was to slow down the *S*s' rate of movement to a level which would be predicted from a model such as Hull's. It should be noted that there is no suggestion of a counterpart to the Crespi depression-effect.

The finding that a low incentive followed by an abrupt change to a high incentive produces *no* decrease in running time does, however, pose a problem. The observed finding could not be explained in Hull's system by a difference in habit strength since Hull's habit strength is a function of $\dot{N}$, the number of trials, and not of the incentive.[8] The explanation might lie in some theoretical construct such as fear, which is commonly assumed to accompany pain as a conditioned mediating reaction. Fear would be assumed some different function of shock than is the incentive, such that their combined effects predict the observed running times.

Summary

Hull's behavior theory predicts that if a change in K, the amount of incentive, be introduced during asymptotic performance, ${}_SE_R$ will rapidly assume a magnitude which would have been obtained had S been con-

[8] Hull, *op. cit.*, 6-7.

sistently exposed to this new value of *K* from the commencement of training. The experiment reported here tested Hull's hypothesis under conditions in which electric shock is the measure of drive, and fixed reduction in shock constitutes the incentive.

All the rats ran an electrified alley 5 ft. long under a 400-v. drive, and were then detained for 15 sec. in the goal-box. One half of the animals were given an immediate 100-v. reduction in shock (low incentive) and the other half an immediate 250-v. reduction in shock (high incentive), for the duration of confinement in the goal-box. After training, the incentives were switched between the two groups of animals. The *S*s for which the incentive decreased showed an increased running time to a magnitude predictable from Hull's theory. The *S*s for which the incentive increased, on the other hand, showed no decrease in running time over 10 trials after the change in incentive had been introduced.

It is hypothesized that the difference between these two outcomes is due to the conditioned mediating response of fear, and not to shock as such.

THE ABSENCE OF PHANTOMS FOR CONGENITALLY MISSING LIMBS

By MARIANNE L. SIMMEL, Brandeis University

In his classical paper on the schema of the body, Pick discussed amputation-phantoms, mentioning in that connection that there are no phantoms of congenitally missing extremities because the absent part has never been represented in the body scheme.[1] This statement seems generally to have been accepted on *a priori* grounds, and it has been reiterated throughout the older literature on phantoms even though no systematic study of such patients had ever been published. Pitres saw one case, as did several other authors.[2] More recently, Browder and Gallagher examined a large group of patients, among them 13 patients with congenital absence of one or more limbs none of whom reported phantom-experiences of extremities missing since birth.[3] The present findings are presented as a further contribution to this problem.[4]

Subjects. The *Ss* of this study were 27 patients with congenitally missing extremities who came to the clinic to have artificial limbs prescribed or adjusted.[5] There were 13 men and boys, and 14 women and girls. Their ages ranged from 6 yr., 4 mo., to 34 yr., with a median age of 10 yr., 9 mo. at the time of the interview. To facilitate the fitting of prostheses, three of the patients had had some surgery of the stump. Table I gives the distribution of congenitally missing limbs in the 27 patients.[6]

* Received for publication July 15, 1960. This report was prepared during the tenure of a U. S. Public Health Service Special Research Fellowship awarded by the National Institute of Mental Health. A supplementary grant from Brandeis University is acknowledged with gratitude.

[1] A. Pick, Zur Pathologie des Bewusstseins vom eigenen Körper; Ein Beitrag aus der Kriegsmedizin, *Neurol. Zentralbl,* 34, 1915, 257-265.

[2] A. Pitres, Etude sur les sensations illusoires des amputés, *Ann. medico-psychol.,* 5, 1897, 5-19.

[3] J. Browder and J. P. Gallagher, Dorsal cordotomy for painful phantom limb, *Ann. Surgery,* 128, 1948, 456-469.

[4] Preliminary results of this study have been discussed in an earlier publication (M. L. Simmel, The conditions of occurrence of phantom limbs, *Proc. Amer. Philos. Soc.,* 102, 1958, 492-500).

[5] Originally there were 28 patients, but one was eliminated by reason of mental defect. All patients were seen from 1956-58 in the bi-weekly Amputee Clinic directed by Dr. Claude N. Lambert of the Department of Orthopedic Surgery, College of Medicine, University of Illinois. The writer is indebted to Dr. Lambert for allowing her to participate in this clinic and to interview his patients.

[6] Casual inquiry made of orthopedic surgeons and one embryologist has failed thus far to yield any explanation of the high incidence of left-below-elbow absence, which usually is attributed to 'chance.' Congenital absence of one or both legs gen-

Procedure. The older patients were interviewed individually. Information about the younger patients was obtained largely by questioning relatives who accompanied the child to the clinic, supplemented by whatever the child himself could contribute. Each interview lasted from a few minutes to half an hour. Some patients were seen on repeated visits, others only once.

The interviews themselves differed markedly from similar interviews with patients who have undergone amputation of a limb. Given the opportunity, most amputees will spontaneously report phantom-experiences. It was to be expected that this would not be true for the patients with congenitally missing limbs. The interviews with these patients began typically with a statement by *E:* "Some patients who have lost an arm or leg have told me that they still feel it. I wonder if you have ever felt your missing arm. Have you ever had any feeling of the hand or the fingers, just like you feel them on the other side?" Most of the patients and their relatives had heard about phantoms, be it from an acquaintance who had lost an extremity, or even from one of the amputee-patients in the waiting room, but many seemed not to have understood what it was all about. Irrespective of *S*'s initial reply, *E* proceeded to ask in detail about feelings of the separate parts, phantom movement, incidents of pain or itching, and so forth. The suggestive character of this detailed inquiry is obvious, but, for the purpose of this study, it seemed safer to err in this rather than in the opposite direction. In any case, as the results below will show, the *S*s were not very suggestible.

TABLE I

DISTRIBUTION OF CONGENITALLY ABSENT EXTREMITIES IN 27 PATIENTS

Position	Fingers, hand	Below elbow	Above elbow	Above knee	Total
Left	2	13	2	0	17
Right	3	3	1	0	7
Bilateral	0	0	2	1	3
Total	5	16	5	1	27

Special problems were, of course, encountered in the interviews with the patients who had ertremities missing bilaterally, insofar as these *S*s lacked a meaningful comparison-object. Patients with unilateral absence of a limb do have the experience of the normal body-part on the other side, and bilateral amputees can retrospectively compare their present experience to what they remember from the time before amputation. No such comparison is available to the person who has lacked both arms or legs from birth.

Results. Of the 27 patients, 25 insisted unequivocally that they had never experienced a phantom of the congenitally missing extremity. As mentioned before, even though many had heard about amputation-phantoms prior to the interview, they had remained puzzled. A considerable number of intelligent and coöperative patients appeared to have diffi-

erally seems to be recognized as occurring much more rarely than that of arms, but again the reason for the difference does not appear to be known.

culty in understanding the questions. A few immediately 'explained' that they had no phantoms because they had never had the limb. Many of the relatives of young patients came forth with the same explanation. These results support Pick's contention that phantoms for congenitally-missing limbs are not experienced by the patients so affected.

Only 2 of the 27 patients reported something other than an unqualified absence of phantoms for the congenitally missing extremity, although this is not to say they gave strong evidence of phantoms. Let the reader judge for himself.

No. 404. Girl. *b.* 7/48, congenital absence of right forearm, below elbow. No surgery. *Initial Interview 11/56.* According to patient and intelligent father, there has never been any evidence of phantom-sensations. Child has never mentioned any feeling of the presence of the missing hand or fingers, has never complained of any pain, itching, or temperature-sensations in the absent parts. No erroneous attempts to grasp objects with right hand ever have been observed. *Second Interview 2/58.* On questioning, *S* says she feels the fingers in the stump, but cannot tell which fingers she feels, cannot tell anything regarding their size, says she does not feel as if she could move them at will. Father does not believe that child feels fingers, thinks she was "giving the right answers." Child has never mentioned anything like that at home. Details essentially as in initial interview.

In the first interview, the child's report agrees with that of the other *S*s in this group. Fifteen months later she apparently reports phantoms, although she is unable to give the usual details. One is inclined to agree with the father's interpretation of the second report and to regard it as due to suggestive questions.

No. 418. Girl, *b.* 9/47. Congenital hypoplasia of right hand. Rudimentary thumb present with some voluntary motion; other fingers absent. Apparently normal sensation in thumb and hand-remnant. X-rays show seemingly normal metacarpal of thumb with a small nubbin of proximal phalanges; complete absence of phalanges of remaining fingers and shortening of residual metacarpals. Carpal bones and distal radius and ulna normal. 8/56 plastic operation to increase useful motion of thumb. *Interview 9/57.* Some voluntary motion of residual bony structures within hand-remnant, especially on its ulnar aspect. During such movement patient says she "feels the little finger, like a real little finger, like the (normal) little finger on the other hand." She is quite definite about this, says it is true only of the little finger, not of the others. Has never experienced any itching, pain, or the like, in this "little finger." At the conclusion of the interview says, "I sometimes imagine the other fingers." No repeat interview. No relative present.

One should note at the outset this child's distinction between "feeling the little finger" and "imagining the other fingers." When adult amputees describe phantoms they report "feeling" the missing part as if it were still present; rarely, if ever, do they use the word "imagine." In fact, one or

two *Ss* have, in the writer's experience, insisted emphatically that they were *not imagining* the missing extremity but that they *felt* it, even though they knew perfectly well that it was gone. It is hardly surprising that a little girl born without the fingers of one hand should at times "imagine" (make believe, day-dream) that she had a normal hand. What may be astonishing is that she so clearly differentiates between sensations and the results of her imagination.

The "feeling of the little finger" on voluntary motion of the residual hand-bones sounds like a report of a *bona fide* phantom. If we accept it as such, we must look for its origin in contemporaneous processes rather than in past experience. Perhaps this phantom is an instance of Feldman's "phantom rod."[7] Feldman applied this term to the illusion created in Hoisington's experiments by putting various pressures and counterpressures on a handle in the hand of the screened *S*.[8] Under those conditions, the *Ss* reported perceiving a rod extending from the handle behind the screen. Feldman attributes this illusion to the perceived motion, *i.e.* the kinesthetic stimulation.

Unfortunately, the motility of the several joints in the patient's hand-remnant was not examined in detail. From the protocols, it appears that voluntary movement was greater at the fifth carpal-metacarpal joint than at the others, and perhaps its excursion was even greater than that in the corresponding joint of the normal hand (due to the absence of tendons and ligaments). The kinesthetic stimulation resulting from such abnormally great excursion may possibly provide a sufficient condition for a 'phantom rod' which, under the special circumstances here, culminates in the experience of a 'phantom finger.'

Summary. Twenty-seven patients with congenitally-missing limbs were questioned regarding phantom-experiences. Twenty-five of the patients reported never to have experienced a phantom. Of the two exceptions, one report was of questionable reliability: the other report may have been of a true phantom based on a contemporaneous kinesthetic illusion. In general, the results support Pick's classical statement that congenitally-missing extremities do not give rise to phantoms.

---

[7] S. Feldman, Phantom limbs, this JOURNAL, 53, 1940, 590-592.

[8] L. B. Hoisington, On the non-visual perception of the length of lifted rods, this JOURNAL, 31, 1920, 114-146.

APPARATUS NOTES

AN IMPROVED CIRCUIT FOR TIME-ON-TARGET AND INTEGRATED ERROR-SCORES IN CONTINUOUS TRACKING

Time-on-target and integrated absolute error are commonly used indices of skill in continuous tracking. Since they frequently measure different aspects of the tracker's behavior, it is often desirable to record both scores. Most circuits presently used to provide either or both of these measures have, however, certain defects which the circuit described here circumvents and moreover makes possible the economical recording of both scores.

Typical of reasonably good, conventional scoring is the circuit employed by Gain and Fitts.[1] A common characteristic of such circuits is the use of diodes in the input-stage to direct a voltage of correct polarity to the next stage. These diodes are inherently non-linear in the low-voltage section of their conductance curves, *i.e.,* output-voltage plotted against input-voltage. This non-linearity causes low-level input signals to be measured incorrectly, since these initially linear signals appear non-linear in the output. Furthermore, diodes do conduct to some extent near absolute zero cross-over, and thus permit leakage-currents which tend to widen the zone where polarity is changing. In other words, when the diode is theoretically not conducting, it does pass unwanted currents owing to leakage. This becomes a serious problem at the extreme of a polarity because, even though the leakage-voltage is small, when it is integrated over time to obtain an integrated absolute error, excessive imprecision results. Measurements of time-on-target are similarly complicated by non-linearity of diodes and by variations in diode-conduction because of fluctuations in temperature.

With DC amplifiers and sensitive relays, a new circuit has been developed which eliminates the disadvantages described above. This circuit is shown in Fig. 1. It is simple, economical, and has proved highly stable and reliable.

The absolute-error section consists of two high-gain amplifiers with a relay in the output which is in series with a diode. The first amplifier is operated 'open loop,' or with maximal gain consistent with stability. Should any low-frequency oscillation or 'motor boating' arise, it is necessary to make a high-gain differentiator by using a 10,000-Ω input and a 20-MΩ feedback shunted by a 0.001-mfd. capacitor.

With high-gain or open-loop operation, polarity-changes are sensed at the input as a function of the gain of the particular amplifier used in the circuit. Since average gains are on the order of 15,000–30,000, the amplifier will reverse output-polarity with a change in polarity of the input for input-voltages on the order of only 0.001 v. The output-stage has a diode to make the relay sensitive to polarity, but it is used only in the linear (high voltage) section of its conduction-curve.

The relay is used to switch in a unity-gain amplifier which operates as a phase-

* This work was supported by a contract between the Office of Naval Research and the Electric Boat Division of General Dynamics Corporation.

[1] Peter Gain and P. W. Fitts, A simplified electronic tracking apparatus, *Wright Air Development Center, Technical Report No. WADC TR-59-44,* August, 1959.

inverter. When polarity reverses, therefore, an absolute value of the input signal is obtained which is highly linear and uncontaminated by diode-leakage. It is now only necessary to direct the output of this stage to a third amplifier, used as an integrator, to obtain an accurate indication of integrated absolute error which may be read out on a good voltmeter.

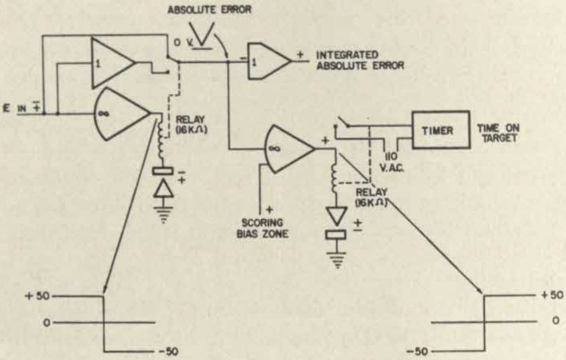


FIG. 1. SCORING CIRCUIT

To obtain time-on-target measures, the signal from the absolute-error curcuit is summed with a reference-bias in another open-loop amplifier. This amplifier acts in the same fashion as the first one, except that the relay in its output switches an electric clutch on a timer which then indicates time within a specified target-zone. The width of this zone is easily varied by adjusting a single potentiometer to change the reference-bias. Normal differences in 'close-open' current of the relays is immaterial, because they are in a circuit where the current is either off or fully on.

Should amplifier-overload occur due to the resistance of the particular relay used, or because the amplifier is incapable of delivering current demand by the resistance of the relay-coil, a series resistor can be added to limit the current through the relay-coil to a safe value. The resistance chosen should be consistent with reliable operation of the relay.

General Dynamics Corporation
Groton, Connecticut

GEORGE SIMONEAU

NOTES AND DISCUSSIONS

CONSTANT ERRORS IN THE MEASUREMENT OF KINESTHETIC FIGURAL AFTER-EFFECTS

Wertheimer reported that during pre-stimulation trials in the test for kinesthetic figural after-effect his right-handed *Ss* overestimated the size of a test-block held between the fingers of the left hand.[1] He attributed his results to the underestimation of the width of the comparison-block held in the right dominant hand and noted that his results were in line with McPherson and Renfrew's finding that, when physically equal, small metal circles are felt simultaneously by the two hands of a blindfolded *S*, the object held in the preferred hand is usually judged smaller.[2]

In an investigation into the effect of meprobamate on the kinesthetic figural after-affect, it was found that on the day on which a placebo was administered a test-block held in the right hand was overestimated during pre-stimulation trials.[3] If Wertheimer's interpretation of his results is correct, there should have been an underestimation of the test-block in the study of the effects of drugs since all the *Ss* were right-handed.

In Wertheimer's own study his left-handed *Ss* also overestimated the test-block held in the left dominant hand and underestimated the width of the comparison-block held in the right non-dominant hand. Wertheimer comments that the number of left-handed *Ss* was too small to provide adequate data to test the effect but the trend in the results is obviously inconsistent with his hypothesis that the important variable responsible for the constant errors is the dominance or non-dominance of the hand.

It was thought that the results in the above studies were more in accord with the hypothesis that during pre-stimulation trials the width of a test-block is overestimated as compared with the width of a comparison-block. Whether or not the dominance or non-dominance of the hand or the function of the block (test or comparison) is the important variable can only be adequately tested by giving the *Ss* trials with the test-block in the left hand and trials with the test-block in the right hand. This procedure was used in the investigation to be reported here.

---

[1] Michael Wertheimer, Constant errors in the measurement of figural after-effects, this JOURNAL, 67, 1954, 543-550.

[2] A. McPherson and S. Renfrew, Asymmetry of perception of size between the right and left hands in normal subjects, *Quart. J. exp. Psychol.*, 5, 1953, 66-74.

[3] C. G. Costello, The effects of meprobamate on kinesthetic figural after-effects, in press.

S was blindfolded throughout the experiment. He was seated at a table on which were placed the test-block (1.5 in. in width) and the comparison-block, which was 0.5 in. wide at the near end, increasing in width to 4 in. at the far end. The *S* was requested to find the place on the comparison-block which felt just as wide as the test-block. Four estimates of the width of the test-block were made, *S* being given alternate ascending (starting from the narrow end) and descending trials (starting at the wide end). After a rest-period of 1 min. the test-block and comparison-block were switched over so as to be held by *S*'s opposite hands, and four more estimates were made. The order of the position of the test-block (to *S*'s right or to the left) was counterbalanced.

A total of 18 *S*s was used. They were all men and all right-handed.

The results for the descending and ascending trials were analyzed separately since it was found in the study of the effects of drugs referred to above that there are important differences between the readings taken by the two methods. Taking the descending trials first, the mean estimate of the width of the test-block held in the right hand was 11.16 (these scores are arbitrary units along the comparison-scale, the point of objective equality to the test-block being 9.0). The difference between this value and physical equality (9.0) is highly significant by a *t*-test ($t = 4.065$, $df. = 17$, $P<0.001$). The mean estimate of the width of the test-block held in the left hand is 11.31. The difference between this value and physical equality is also highly significant ($t = 6.251$, $df. = 17$, $P<0.001$). The difference between the mean estimate when the test-block is in the right hand (11.16) is not significantly different from the mean estimate when the test-block is held in the left hand (11.31).

The above results are clearly inconsistent with Wertheimer's hypothesis that the differences he found are related to handedness and are consistent with the writer's hypothesis that they are related to the function of the block.

Wertheimer used both ascending and descending trials but did not analyze them separately.[4] In the present study, with readings taken on ascending trials, the mean estimate of the width of the test-block, when it is held in the right hand, is 8.175. The mean estimate when held in the left hand is 8.178. Neither of these mean estimates is significantly different from physical equality and they are not significantly different from one another. It would appear, then, that the constant errors that occur during pre-stimulation trials in the test for kinesthetic figural after-effect occur only on descending trials, *i.e.* when *S* is moving his hand from the wide end of the comparison-block to the narrow end.

Institute of Psychiatry
University of London

CHARLES G. COSTELLO

---

[4] Wertheimer, Personal communication.

A NOTE ON 'ATTITUDE AND DISTANCE-ESTIMATION AS VARIABLES IN SIZE-MATCHING'

Jenkin and Hyman, in a paper published in this JOURNAL, describe an experiment in which *S*s made a series of size-matches with analytic and with objective instructions.[1] For example, an adjustable comparison-triangle 30 ft. from *S* was matched with a standard triangle 15 ft. away. The *S*s also estimated the distance of the mounting-board when 30 ft. away. The inter-*S* correlations between these size-matches and distance-estimates are interpreted by the authors as reflecting the relationship between judged size and estimated distance. Size-match is not, however, directly equivalent to judged size: this difficulty is discussed below, and an explanation of the correlations is suggested.

Size-matches indicate *relative* perceived size. It would, therefore, have been appropriate to relate the size-matches to relative distance-judgments, as did Gruber.[2] Alternatively, the perceived size of the nearer object may be assumed equal for all *S*s. It then follows that the size-match settings of the comparison-triangle are such that its judged size is also equal for all *S*s. From this it may be inferred that the median setting of the comparison-triangle would be judged relatively *small* by an *S* who gave a *large* size-match, and large by one giving a small size-match. In this sense, size-matches are the 'reciprocal' of judged size.

Thus, the direction of the correlation between size-match and estimated distance is opposite to that between judged size and estimated distance. This relationship is not clearly stated in Jenkin and Hyman's paper, and the signs of the correlations are ambiguous. Jenkin elsewhere explains that the signs given in the paper refer to implicit judged size and not size-match.[3] Therefore the statement on p. 76 of the report should read: "Hence individuals who make a judgment without attending to projective size and who form an impression of relative 'farness' tend also to respond with a size-judgment [not *size-match*] which is relatively small."

Now if the interpretation of size-match in terms of judged size is accepted, Jenkin and Hyman's results are important. For their size-distance correlations are non-linear and negative in direction with objective instructions, while a linear positive correlation is predicted by the commonly accepted hypothesis of invariance.[4] A small positive correlation

---

[1] Noël Jenkin and Ray Hyman, Attitude and distance-estimation as variables in size-matching, this JOURNAL, 72, 1959, 68-76.

[2] H. E. Gruber, The relation of perceived size to perceived distance, this JOURNAL, 67, 1954, 411-26.

[3] In correspondence with the writer.

[4] W. S. Duke-Elder, *Textbook of Ophthalmology*, 1, 1932, 1071; Harold Schlosberg, A note on depth perception, size constancy, and related topics, *Psychol. Rev.*,

occurs only with analytic instructions. These results are supported by Gilinsky's observation that perceived size varies with distance in the manner predicted by her formula (partially derived from the hypothesis of invariance) only with analytic instructions.[5] Similarly, Weckowicz has shown that schizophrenic patients, whose size-judgments are nearer the size of the visual-angle than normal *Ss*', give a small positive correlation between 'size-constancy' and 'distance-constancy.'[6]

An experiment by the writer suggests that Jenkin and Hyman's correlations result from *Ss*' inferences, which in turn depend on their beliefs regarding the nature of the stimuli.[7] If two stimulus-objects are considered to be of the same physical size, a negative size-distance association tends to result; otherwise, a positive association is found. These beliefs would presumably be held by *Ss* given objective instructions and analytic instructions respectively. Thus the positive and negative correlations in Jenkin and Hyman's study may have arisen. It is interesting that Jenkin, too, concludes for quite different reasons that the way in which an observer 'identifies' even an unfamiliar stimulus-object may affect its judged size.[8]

If the above suggestions are correct, Jenkin and Hyman's data are significant in two respects. Further evidence is provided that the prediction of a positive linear correlation between perceived size and perceived distance does not hold for inter-*S* correlations with constant stimulus-conditions; and that an observer's inferences may influence the size-distance relationship in unrestricted as well as reduced-cue conditions.

Manchester University, England E. E. RUMP

THIRTY-THIRD ANNUAL MEETING OF THE MID-WESTERN PSYCHOLOGICAL ASSOCIATION

The Midwestern Psychological Association met at the Hotel Morrison, Chicago, on May 4, 5, and 6, 1961. George S. Speer, Illinois Institute of Technology, served as Convention Manager and chairman of the Local Arrangements Committee. The number of persons registered was 1,801.

57, 1950, 314-317; T. E. Weckowicz, Robert Sommer, and R. W. Hall, Distance constancy in schizophrenic patients, *J. Ment. Sci.*, 104, 1958, 1174-1182; R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 3rd ed., 1955, 462; 484.

[5] A. S. Gilinsky, The effect of attitude upon the perception of size, this JOURNAL, 68, 1955, 173-192.

[6] Weckowicz *et al.*, *op. cit.*,

[7] E. E. Rump, The relationship between perceived size and perceived distance, *Brit. J. Psychol.*, 52, 1961, 111-124.

[8] Jenkin, A relationship between increments of distance and estimates of objective size, this JOURNAL, 72, 1959, 345-363.

The number of members newly elected or reinstated during the year was 258, bringing the current membership to 2,461.

The program, planned by a committee under the chairmanship of Willard A. Kerr, Illinois Institute of Technology, consisted of 187 volunteered papers presented in 27 sessions, 6 symposia, a workshop, and a program sponsored by Psi Chi which included 5 volunteered papers and an invited address.

The paper reading sessions were concerned with: social psychology; psychopharmacology; verbal learning; perception; animal reinforcement conditions; personality assessment; higher processes; sensory processes; mathematical behavior models; brain functions; verbal behavior; animal learning; personality dynamics; comparative and genetic psychology; somatic variables and learning; schizophrenia; attitudes and attitude change; human conditioning; engineering, personnel and training; animal motivation; early experience; psychodiagnostic testing; human motive-incentive conditions; and clinical psychology.

Symposia and open meetings were concerned with: models and organic correlates of psychological processes; psychological inferences concerning organic brain damage; workshop for State Associations by APA Board of Professional Affairs; concepts in the experimental study of human motivation; current research on experimental gastric ulcer; action training techniques in industry; and the role of the internship in predoctoral clinical training.

Marion E. Bunch, Washington University, delivered the Presidential Address, "Experimental extinction in learning and memory."

I. E. Farber, University of Iowa, was elected President, 1961-62, and Donald R. Meyer, Ohio State University, was elected Secretary-Treasurer, 1961-64. Elected to the Council were: Edward L. Walker, University of Michigan, 1961-64; Ernest J. McCormick, Purdue University, 1961-63; and James M. Vanderplas, Washington University, 1961-62.

The 1962 meeting will be held May 3, 4, and 5 at the Hotel Morrison, Chicago.

I. E. FARBER

State University of Iowa

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

The Psychology of Perception. By WILLIAM N. DEMBER. New York, Henry Holt and Company, 1960. Pp. xi, 402. $6.50.

Due to changing conceptual emphasis, perception has become a difficult area to delimit. Dember has written a text for an introductory college course in visual perception whose content ranges from a study of the simple psychophysical relationships formerly considered under "sensation" to a study of psychological processes highly motivational and cognitive in character. The author states that the purposes of the book are to acquaint the reader with the "principles that have emerged from the empirical investigation of perception" (p. 10) and "to expose problems in the hope that they will help stimulate future thinking and research" (p. 11). The book is written in a lucid straightforward style and lives up to the author's aims in its discussion of particular topics. Due to an inadequate coverage of many of the major subjects in the field, and its failure to meet fully the basic problems of definition and approach posed by the subject material, the book does not however fulfill the author's aims in a wider sense.

The topics covered and the chapters allotted to them are: three chapters discuss problems connected with psychophysical methods and their application to threshold-measurement and stimulus-scaling; one chapter examines some of the simpler relationships of visual psychophysics; two chapters are concerned with the organization and interaction of patterns of stimulation and examine such relational phenomena as visual organization, spatial perception, and sensory-motor interaction; one chapter examines the effects of learning on perception; two chapters review the findings concerning the effects of set and motivation on perception; and a final chapter looks at the act of perceiving from the point of view of being gratifying in its own right. It is in its discussion of selected topics in measurement, visual psychophysics, the effects of learning, and the motivation of perception that the book makes its most novel contribution as a text. Here Dember makes available to the undergraduate student an introduction to the work of Coombs, to that of Blackwell and his associates, to the studies on early and extra experience, and to studies on the reinforcing properties of perception.

The book is least satisfying in its discussion of the classical problems of perception set by Koffka's question "Why do things look as they do?" A rewarding section of any introductory course is the discussion of the attributive and functional differences between the perception of color as a characteristic of light and color as a characteristic of surface. The simple notions brought from physics and too often from elementary psychology courses are here dispelled and the student introduced to the complex problems connected with the perception of surface and color. The perception of color is, only briefly discussed by Dember in connection with contrast, memory-color, and its disappearances in the *Ganzfeld.* Though Dember expresses his concern with the informational aspects of stimulation, his considera-

tions seem to have been primarily restricted to detection-theory. In discussing visual organization, the book simply lists and illustrates the Gestalt laws of organization. The recent studies of Hochberg and by Attneave in which the often equivocal Gestalt laws are formulated more precisely by viewing the organizational process as reflecting an economical description of the stimulus input are not mentioned. By failing to grapple with the problems of defining the stimulus, the perceptual response, and the nature of intervening mechanisms, the book also fails to acquaint the reader with the different general approachs to perception. In the discussion of space perception, for example, the contrasting views of Gibson and Brunswick concerning the perceptual process underlying an organism's relationship to its spatial environment are not introduced. The narrow approach of the book to the traditional problems of perception results in other major perceptual topics being inadequately treated, such as the perception of motion and the role of learning in the development of constancy.

In summary, the value of this book is in making available to students a highly readable and superior introduction to selected problems. The failure, however, to treat adequately what to the reviewer are some of the central problems of perception strongly restricts its usefulness as a general text.

JACOB BECK

University of Pennsylvania

Visual Field Defects after Penetrating Missile Wounds of the Brain. By HANS-LUKAS TEUBER, WILLIAM S. BATTERSBY, and MORRIS B. BENDER. Cambridge, Mass., published for The Commonwealth Fund by Harvard University Press, 1960. Pp. xi, 143, $4.75.

This monograph is a summary of one portion of the results of a continuing research program on the after-effects of brain injury in man. In undertaking these studies, the authors have implicitly or explicitly set themselves three distinct tasks. The first task is that of developing methods for the detailed and reliable analysis of deficits in visual function. The second is the accumulation of a group of subjects and a body of data large enough and sufficiently homogeneous so that generalizations may be made with some confidence. Finally, there is the task of relating the obtained results to the existing body of knowledge and theory as to the role of the geniculo-striate system in visual function. The author's have succeeded in all three tasks.

These studies derive much of their value from the investigators' basic assumption that "abnormal function after brain injury can and should be studied by methods drawn from psychophysics and comparative psychology." They are further distinguished by the care and ingenuity with which these methods are employed. The methodological considerations are important ones because of the limitations of clinical observation as a tool for the analysis of brain function. Thus, the primary measures of visual field defects involve the perimetric and campimetric plotting of central and peripheral fields, and there has been much criticism of the reliability of these measures. In view of the fact that the perimetric data occupy almost the entire first half of the book, it is gratifying to note the authors' concern with the methodological problems involved, *e.g.* retest reliability, control of eye movements, etc. It is even more gratifying to note, in addition to the perimetric studies, the inclusion of tests of flicker perimetry, dark adaptation, depth and motion perception and a number of tests of a more general perceptual nature.

The cases studied were drawn from a population of "previously healthy young adults." The absence of complications arising from neoplastic diseases or vascular disturbances is an advantage, as is the fact that both acute and chronic cases were available for comparison.

Two complementary bodies of data are presented. The first was obtained from the perimetric studies of field defects and is the basis for an exhaustive discussion of what James might have called "the varieties of scotomatous experiences." The perimetric data are carefully analyzed with a view toward determining the extent to which the visual field defects conformed to existing concepts of topographical projection in the visual system. A number of revisions of these concepts are suggested.

The second group of results comes from the study of what the authors term "functioning versus plotted fields." These data indicates that the picture of visual deficits provided by perimetric data is an incomplete one. On the one hand, there appear to be various compensatory processes which make the "patients' functioning fields seem wider than the perimetric fields." On the other hand, there are "subtle but significant changes in other visual functions, even in those parts of the visual field which seemed intact according to the perimetric plot." These data are the basis for perceptive discussions of the distinction between the projection of visual fields and the representation of various levels of visual function. There is a brief, though stimulating, discussion of the problem of the "agnosias" and the concept of the dissociation of "higher" and "lower" visual function. The interpretations throughout are advanced with caution, tempered as they are by the authors' recognition of the one inevitable limitation of such studies—the absence of precise information as to the location and extent of the lesion.

In the preface to this volume, Teuber acknowledges the group's indebtedness to the tradition exemplified by the work of Goldstein, Lashley, and Klüver. The present monograph is a significant contribution to that tradition.

Laboratory of Neurophysiology — H. PHILIP ZEIGLER
University of Wisconsin

Physics and Medicine of the Atmosphere and Space. Edited by OTIS O. BENSON, JR., and HUBERTUS STRUGHOLD. John Wiley and Sons, Inc., 1960. Pp. xviii, 645. $12.50.

This book represents the proceedings of the Second International Symposium on the Physics and Medicine of the Atmosphere and Space held in 1958. The symposium was sponsored by the U.S. Air Force School of Aviation Medicine and was held in San Antonio, Texas.

This conference followed the first Symposium by seven years. The papers of the first meeting were also published in book form (*Physics and Medicine of the Upper Atmosphere,* edited by Clayton S. White and Otis O. Benson, Jr., The University of New Mexico Press, 1952.) There are several elements of contrast between the contents of the two sessions. These reflect the advances in aeronautical and space vehicle capability achieved during that period of time. Also, they represent the current anticipatory thinking on man's role in exploration of extraterrestial environments. Thus, where the First Symposium dealt with the biological significance of the earth's upper atmosphere, much of the present work is dedicated to the interplane-

tary medium. Except when maneuvering near the earth's surface, the space vehicle is not subject to aerodynamic forces but to the laws of celestial mechanics. The velocities considered are those required to achieve insertion into orbit or escape. Consideration of the circumstances and "hazards" of space travel accounts for most of the Symposium.

At the same time, the contents of the Symposium reflect the cross-discipline nature of the subject of Space Medicine. The broad aims of this field have come to be identified as divided between the preventive medical aspects of manned space operations and astrobiology—the study of life on other celestial bodies. The book contains two chapters on the knowledge of the physical environments of the moon and the planets and a single chapter concerning experimental work in astrobiology. The remainder of the Symposium deals with factors which are important in the design of space "hardware" for manned operations and in the selection and training of the human operator. The advice of astrophysicists, physicians, physiologists, aeronautical engineers, and psychologists has been enlisted in this effort. The following paragraphs deal with those sections with treat psychological problems in particular.

Three chapters deal with the subject of nutrition of a human occupant during a space flight. Hans Clamann outlines the problems in human metabolism which may be expected in a closed ecological system. Robert Tischer and Jack Meyers discuss methods of providing nutrition during space voyages. However, there is no real discussion here of psychological aspects of space feeding.

Psychological problems concerned with man's existence in a closed ecological system are treated by a group of authors. George Steinkamp describes the sealed cabin simulator used at the U.S. Air Force School of Aviation Medicine. George Hauty, one of his collaborators in this set of experiments, outlines three problem areas and describes the experimental results which his group obtained. He discusses behavior during sensory deprivation, efficiency of task performance with changes in 24 hour work-rest cycles, and the general subject of skill fatigue. He suggests that this latter aspect may impose one of the greatest limitations on man's useful functioning in space vehicles. From the psychologist's point of view, this chapter is made interesting by his suggestions as to possible direction for future investigation.

Heinrich Rose attempts some predictions as to what the human operator in a space vehicle may perceive of other objects in space and on planets beneath their atmospheres. He attempts to illustrate the great limitations imposed by the light intensities and contrasts, the velocities of the objects to be perceived and their sizes in terms of visual angles subtended. Gerathewohl and Ward's chapter on weightlessness serves to illustrate how little is really known of this subject. They have listed the results of the few studies performed. The chapter entitled "Biomedical Aspects of Orbital Flight" serves to illustrate the wide variety of factors which must be anticipated, depending on the nature of the mission planned. This is a general treatment of the subject covering a wide area of consideration. Finally, the speculative treatment of time dilation in vehicles which achieve relativistic velocities by Evan Goltra is interesting as it anticipates greater vehicle capability at some future date.

Ross A. McFarland

Harvard University

Rorschach Psychology. Edited by MARIA A. RICKERS-OVSIANKINA. New York, John Wiley and Sons, Inc., 1960. Pp. xvi, 483. $8.50.

This book of fourteen chapters, contributed by seventeen authors, will be of interest and value to the reasonably well advanced student of the Rorschach test. It is not intended, presumably, for the novice. Nor is it a volume for the experimentalist or other psychologist who has not had fairly extensive experience in the use of the Rorschach with a variety of subjects. Anyone who is to evaluate Rorschach psychology, as discussed in this volume, should have had practical experience with the test.

The book, on the whole, is a healthy effort to relate Rorschach responses and their interpretation to some theoretical concepts such as figure-ground, *Prägnanz,* changes in perceptions under varied physiological conditions (*e.g.* anoxia). An effort is also made to relate Rorschach findings to motivational and general personality theory, and to developmental aspects of personality.

The several authors make more or less valuable contributions to these and other topics; but they do not and cannot be expected to exploit them to their fullest within the scope of a single volume. For example, the development aspects of color responses are presented, but no attention is given to the psychophysical and the psychophysiological problems of color vision. Too little attention is given to cultural factors and differences, as they affect responses; very little is said about the significance of differences in intelligence, as such. And, there is little systematic discussion of Rorschach responses as related to general theory of personality.

The editor of the volume states that "Rorschach's chief working principles" are to be examined "in the light of contemporary psychological knowledge." It appears to me that this promise is not adequately fulfilled throughout the book, and that in too many of the chapters there is much more emphasis upon techniques than upon "psychological knowledge," by which I assume theory and principles are meant.

These deficiencies, however, are indications only of the fact that the editor and the authors had to make their choice of topics; and those omitted probably seemed to be among the less significant ones for their purposes. The absence of treatment of certain topics some of us would like to have discussed does not detract from the considerable value of that material which is presented.

As is the case with all multi-authored books, the chapters are not of equal merit or value. The introduction is followed by nine chapters on categories of analysis, two on the test-pattern, one on reliability, and one on validity. There is also an appendix giving a tabular comparison of scoring systems. In my opinion, the best written and generally most valuable chapters are these: "The Experience Type," by J. L. Singer; "Hermann Rorschach and Personality Research," by Lois and Gardner Murphy; "Reliability Re-examined," by J. D. Holzberg. The chapter which, it appears to me, contributes least is the one on "Validity: The Search for a Constant in a Universe of Variables," by J. G. Harris, Jr. The writing in this chapter is 'precious,' strives for 'literary effect,' makes a number of irrelevant historical references, obfuscates by introducing unnecessary jargon, and tires this reviewer with its figurative language (*e.g.* "agents of justice," "circumferential touches of global logic," "ameboid movements of the ink-blot specialists," "exhortations of reformers"—all on one page—"the shadow of appeals," "keys to the scientific kingdom," "sacrificial commitment of the interpreter)." It is regrettable that validity, the most important

problem of Rorschach or in any other type of testing, has been dealt with as it has in this chapter, whereas clarification and lucidity are demanded.

In their chapter, the Murphys write: "It is not because Rorschach devised a test that his position will remain revered [*sic*] in psychology; it is because his test is the embodiment of a sensitive, exploratory, original view of human individuality, unfinished but capable nevertheless of being measured, the measures interrelated, moment by moment, as the demands of life change. It is because the Rorschach test is the expression of a profound conception of personality that it is important in psychology today" (p. 347). It would have been fortunate if Professor Rickers-Ovsiankina had obtained this statement from the Murphys, to begin with, and had given it to the other contributors, to be used by them as the basic theme of the book. We might, then, have had a series of chapters of more nearly equal significance with regard to Rorschach psychology. As it is, however, this volume can be read with profit by those concerned with the utilization of the instrument, with research problems it has created, and with its probable contributions to psychology as a behavioral science and a clinical profession.

FRANK S. FREEMAN

Cornell University

Stuttering and Personality Dynamics: Play Therapy, Projection Therapy, and Counseling. By ALBERT T. MURPHY and RUTH M. FITZSIMONS. New York, The Ronald Press Company, 1960. Pp. vii, 519, 13 pl. $6.50.

This book is a series of essays on various phases of stuttering. Each essay sets its own context. Very few sentences of any chapter of this book can be meaningfully quoted alone. "Stuttering behavior is primarily a psychogenically motivated symptom which manifests itself most discernibly in oral functions" (p. 17), say the authors. The meaning of the term *oral* is left to context.

Similarly context is required to give meaning to the statement: "Stuttering is one pattern of adjustment. It is the most effective adjusment possible for the individual at a given moment. It is deeply purposive" (p. 19). The same could be said about learning to engage in a perseverative game called patty cake. "Stuttering can be defined [sic] as 'what a person is' " (p. 172). This sentence means nothing without context.

Even when the authors announce in italics that they are about to give us a soul-satisfying definition they say: "Stuttering is a learned, nonintegrative, selfdefensive reaction to anxiety or fear of threatening circumstances with which the person feels incapable of coping" (p. 145). So are many cases of alcoholism. Apparently the authors wrote this book in the process of a series of conferences with each other, on a basis of client-centered counseling, in which the clients did a lot of non-structured free associating. Putting together however, all of the fragmentary statements about stuttering, and gleaning from context a meaning for each, the reader gets the notion that the authors believe that stuttering is a disturbance of oral communication caused by an uncomfortable social adjustment of the speaker to his auditor. "A constitutional factor in stuttering *may* exist. . . . But beyond this verbal tip of the cap to the possibility of a 'somatic variant,' this book devotes it pages exclusively to the treatment of stuttering phenomena from a clearly psychodynamic point of view" (p. 45). This means, in short: *if* or *when*

stuttering is caused by morbid psychodynamics, then the cure of stuttering lies in correcting the dynamics. It means this, and no more. It means that if or when the automobile engine misses or stutters, due to stoppage of the fuel line, the solution of the problem is to remove the stoppage. "Clearly, therapy must be structured [sic] so as to cope with the range of verbal and nonverbal dynamics discussed, and although emphasis must be on speech therapy which is psychodynamically based and propelled, nevertheless there is a continuum of therapy procedures just as there is a personality-speech continuum of unique stuttering persons, many of whom will profit by more directive therapies than are considered in this book but which are available elsewhere" (p. 173). The reader of *Stuttering and Personality Dynamics* will therefore, look to Murphy and FitzSimons for help with the management of stutterers, only *if* he agrees with their statement in the preface: "We are convinced that stuttering speech is a symptom of deep-seated personal difficulties" (p. v).

Even if the reader accepts this thesis, he may still find it difficult to draw from this book real, solid, practical help. The vagueness of exposition, alluded to above, may confuse the reader. A reflection of this vagueness is to be found in the index. *Psychodynamics,* a word used in the text, the heart of the discussion of stuttering, does not even appear in the index. The index is comprehensive enough, however, to include 18 columns of items, including *anal phase of development, oral phase of development, castration anxiety, penis envy, conscience,* and *guess what!;* but it omits the last word in the book's main title, *dynamics.* The book is clearly for those who are already sophisticates in the philosophy of Carl Rogers, and for those skilled in non-directive therapy. Others may find the book too vague and esoteric to be of much help with stuttering.

Brooklyn College ROBERT WEST

Tactics of Scientific Research: Evaluating Experimental Data in Psychology. By MURRAY SIDMAN. New York, Basic Books, Inc., 1960. Pp. x, 428. $7.50.

This book describes the method of research that has developed in the study of operant conditioning during the past few decades. The experimental examples are taken entirely from the area of animal psychology. It is clear, however, that Sidman's intention is to convince the reader that this method is the paradigm of research in all fields of experimental psychology.

The central theme is that a lawful functional relation demonstrated in behavior is significant if found only in one subject and at a brief span of time. When a Newton demonstrates the law of gravitation with one metal ball, the importance of the experiment does not depend on numerous measures of the motion of that same ball, of many different balls, of different metals, and of variously shaped objects, etc., nor is confidence in the conclusions to be coaxed from any statistical manipulation of the data. When the psychologist attempts to deal with variability by resorting to statistical analysis of group-data, he usually loses sight of the basic phenomenon which exists only in the behavior of individuals. The real spadework of research in psychology, as in other sciences, is controlling the sources of variability.

Sidman subjects many "sacred cows" of present-day experimental psychology to ruthless criticism. The reviewer believes that this is quite timely, as psychology seems threatened with becoming lost in the wilderness of practical problems where it seems to be getting further and further away from the goal of scientific respect-

ability. The following is a partial listing of conventions criticized: (1) The use of separate group-means to establish a relation between behavior and an independent variable; (2) the stress placed on hypothesis or theory in the evaluation of data; (3) the preliminary nature of an 'exploratory' investigation; (4) the uncritical acceptance of the concept of intrinsic variability; and (5) the idea that logicians have developed a satisfactory set of rules for the designing of experiments.

Considered in the book are the following topics: The basis for judging the importance of scientific data; problems of reliability and generality; replication; the nature of variability; and specific experimental designs. The designs and techniques described are pointed specifically at problems of the behavioral processes in individual subjects, such as, the achieving of steady states, the handling of transitional states, the establishing of behavior baselines and techniques of control. The author makes it clear that problems of experimental design cannot be separated from the nature of the data being studied. His presentation is specifically oriented toward the area of operant conditioning, in which he has been one of the productive experimenters. Little effort is made to develop analogues in other areas of psychological research. This is left as an implicit challenge to the readers who are working actively in other areas.

Before he has read to the end of the first chapter, the reader will realize that this book does not belong with his other volumes on experimental design. It is primarily concerned with controlling the data and handling of the problems of research rather than with abstract mathematical and statistical procedures for extracting generalizations from poorly controlled data. Sidman does not lay down a convenient set of rules which the student may follow blindly. He specifically denies that the logicians have as yet worked out a rationale for what productive psychologists do.

Sidman has given us, not just another treatise on experimental design, but also a review of the results, with emphasis on their significance of much of the recent work on operant conditioning. The students of psychology who have not systematically kept up with the fairly steady flow of reports should be grateful for this exposition; and those who read it will most likely be profoundly impressed with the prospect it unfolds.

HENRY N. PETERS

VA Hospital, Jefferson Barracks, Mo.

The Strategy of Conflict. By THOMAS C. SCHELLING. Cambridge, Harvard University Press, 1960. Pp. vii, 309. $6.25.

Observing that "the military services, in contrast to almost any other sizable and respectable profession, have no identifiable academic counterpart" (p. 8), Schelling suggests that the theory of games, broadly interpreted and extended in a variety of ways, may serve as the nucleus of such a scientific discipline. His book is composed of ten loosely interlocked, slightly redundant essays, several of which have previously appeared as journal articles, and three appendices. His concerns are theories of bargaining, threats, deterrence, tacit and explicit communication, and the like, and applications to real conflicts, especially those having geopolitical and military significance.

Schelling is perceptively critical of current formulations of game-theory, charging (correctly, I believe) that its focus on the limiting case of two perfectly opposed protagonists (zero-sum games) has led to misconceptions about the significant

features of the much more common and important cases of mixed (or partial) conflict. A number of failings are discussed in detail and are liberally illustrated by examples that range from the 'warfare' of discipline between parent and child, through the tacit and explicit agreements and use of threats inherent in driving an automobile, to the possibly parallel and much more serious notions involved in having and threatening to use military power. Just how parallel they are—to what extent governments and individuals can be usefully considered as theoretical entities of the same sort—is often the crux in evaluating the significance of his examples.

To the psychologist, Schelling's demonstrations that factors not now captured in the formal theories are important and sometimes controlling features of behavior suggests a range of researchable problems not evident from the classical theory, and he reports results from a number of informal studies. These range from asking pairs of *S*s to name heads or tails independently, agreement resulting in a prize 36 of 42 *S*s chose heads), to extremely rich mixed conflict games (see the example discussed on pp. 102f.). What is not provided, however, is an adequate theoretical framework about which to organize and interpret such empirical work. His contribution is cogent criticism and indications of what is needed, but very little is really added to the theory. So, for a time, we may see developing what could easily become a sprawling empirical literature of isolated and special findings.

To the citizen, the most fascinating feature of the book is Schelling's analyses of such problems as limited war, surprise attack, and arms control. Here Schelling is in his element, seriously considering serious problems and, through a shrewd use of verbal reasoning, arriving at interesting and often surprising insights and conclusions; and here is where the reader must be most wary. Following the traditions of military, economic, and political discussions, but incorporating recent insights into questions of strategy, these analyses treat in pathetically simple terms problems of enormous complexity and of, literally, human existence. This is no criticism of Schelling, for we have no better tools and so must use those we have, but the need is great to ferret out implicit assumptions and values, to question the simplifications made, to wonder about incomplete consideration of possibilities, and to be tuned for logical imperfections. Most such discussions of international conflict, including some of Schelling's, seem to be infused with a form of rationalistic paranoia. Similarly, spare models approached with the same pristine logic seem to suggest that it is impossible for societies of largely unarmed people to form and be stable as well as to recommend that the best way to prevent the feared world holocaust may be to indulge in arms races and have high levels of arms. One wonders!

University of Pennsylvania R. DUNCAN LUCE

Visual Space Perception. By W. H. ITTELSON. New York, Springer Publishing Co., 1960, Pp. ix, 212. $6.00.

Ittelson presents what he designates as a "wide ranging" theory of perception in which some data from space perception are incidental but handy for illustration. The core of Ittelson's transactionalism is the unconscious which creates the abstraction which he defines as perceiving, with the end product being, "the creation of certainty out of uncertainty or probability" (p. 38).

Ittelson proceeds to arm the unconscious with an "invariance hypothesis" as a

shield and with a sheaf of darts which he calls the visual space cues. With these weapons he marches through five chapters of hand picked data and numbers and emerges into special problems of perception, unscathed, without changing an issue, and with hardly the loss of a cue.

He alleges that, "specifically, the invariance hypothesis states that with impingement invariant, perception must necessarily be of one of the equivalent configurations defined by that impingement" (p. 53). By impingement he means the light to one or both eyes and he leaves it to the reader to define "invariant." This strong position is immediately vitiated by his own qualms when he states, "It is never safe to assume arbitrarily that any given perception falls within the limits of the invariance hypothesis" (p. 53). Ittelson's other weapon is the sheaf of cues found in any elementary textbook on psychology. He defends his choice by saying that they have basic validity and girds himself thus: "It seems unnecessary and, in fact, undesirable to approach the study of space perception by any other road" (p. 43). In his chapters on cues of visual space he goes over each one in turn and finally realizes his difficulties. "Cues enter into any concrete transaction in combination. An understanding of the visual space cues necessarily includes a study of these combinations. Taken in its greatest generality, this is, at our present state of knowledge, an impossibly complex problem" (p. 51).

Ittelson wraps up space perception in the heading, "From WHEATSTONE to AMES" (p. 111). He is obviously a disciple of Ames and dedicates his book to him. As might be expected, the principal questions of the disciple are those of his master: (1) "Given a particular impingement, what are the related physical configurations?" (p. 48); (2) "Given a perception what is the related physical world?" (p. 113); (3) "Given a particular disparity, what is the accompanying arrangement?" (p. 113). These are venerable questions and it is easy to understand how Ames became fascinated by them. These questions in conjunction with his observations while wearing aniseikonic glasses led Ames to the construction of his now famous illusion of the distorted room (etc.), Ittelson cloaks the study of these illusions in the dogma of transactionalism, with no predictive, functional interrelationships between the dogma and the illusions.

Ittelson's scholarly housekeeping in *Visual Space Perception* leaves much to be desired. The book *is not* a survey of the literature. There is *no debate* of the various theories of perception. The author chooses only the literature which he feels appropriate to his cause. None of his quotations are paginated. Repeatedly he cites multiple authors in support of his arguments without specifying which of the multiple works of individual authors he has in mind. To make matters worse he names authors and fails to include their works in his reference list. He repeatedly fails to cite the literature on which he bases many of his conclusions. An advanced graduate student would have difficulty in reading this book intelligently. In conclusion, it seems fair to comment that publishers seldom publish term papers and professors seldom write term papers. This book is the unhappy consequence of such a coincidence.

Olin W. Smith
Cornell University

Decisions, Values, and Groups. Edited by Dorothy Willner. New York, Pergamon Press, 1960. Pp. xxix, 348. $12.50.

This volume contains twenty-three reports from the First Interdisciplinary Con-

ference in the Behavioral Science Division which was held at the University of New Mexico under the sponsorship of the Air Force Office of Scientific Research. In its Introduction, Anatol Rapaport presents the point of view that the three areas of research mentioned in the title of this book (decisions, values, and groups) are currently studied by people in a number of different academic disciplines, and therefore can be said to be interdisciplinary in the nature of their content. Actually, the book itself is divided into five sections which are not perfectly correlated with the three areas mentioned in its title.

Section I is entitled "Mathematical models in decision processes." Here are collected papers by Edwards, indicating some new directions of research on decision processes; by Brodbeck, discussing the application of mathematical models to the social sciences; by Scodel, Ratoosh, and Minas, reporting experimental results on some personality correlates of decision making under conditions of risk; by Rapaport, Hays, and Birch, reporting findings in a study of individual styles of hypothesis-formation; and by Chipman, discussing a model of stochastic choice and subjective probability. Section II has the title, "Conceptualizations and designs for research in values and evaluative processes." It begins with a paper by Travers, who proposes increasing use of the psychological laboratory in the study of individual differences. This is followed by reports by McCord and McCord, on the structure of conscience; by Wertheimer, describing an approach to the study of "person cognition;" by Scott, who reports on some relationships between group interaction and personal values; and by Rose who describes briefly some findings on the organization of microcultures occurring in small groups. Part III is entitled "Theoretical contributions to small group research." Beginning with a discussion of methodology by Borgatta, this section continues with a paper discussing the relationship between certain properties of biological systems and aggregates of persons by Campbell; a report of a study on seeking or avoiding self-evaluation by individuals in groups by Willerman, Lewit, and Tellegen; an experimental study of message communication by MacLay and Newman; a study of measuring individual prominence in groups, by Shaw; and some findings on the effects of social support on the influence of propaganda, by Brodbeck. Part IV concerns "Psychodynamic patterns of behavior." In this part, Gerard discusses some physiological parameters of the nervous system as determiners of behavior; Worchel presents a theory and experimental study of hostility; Charms and Rosenbaum discuss the effects of vicarious experience in modifying behavior; Westley and Epstein describe methods for measuring the psycho-social organization of the family; and Zimmer and Foy provide a systematic account of depression and its determinants. A final section of the book, Part IV, is entitled "Special military problems." Here one finds an article by Dailey concerning an analysis of motivation as revealed by reënlistment rates; and one by Campbell summarizing physiological and psychological factors affecting manned space flight.

Generally speaking, the ideas presented in this volume represent new directions of thought in psychology, and particularly in social psychology. The interested reader may expect to be stimulated to consider some new conceptualizations, rather than to find discussions which fit currently traditional molds. While these statements cannot apply to every one of the reports in the volume, they are applicable to most of them. Although there is little evidence here of the cross-fertilization

that is sometimes claimed for the "interdisciplinary" approach, these reports evidence much originality.

Princeton University ROBERT M. GAGNÉ

Guidance of the Young Child. By LOUISE M. LANGFORD. New York, John Wiley and Sons, 1960. Pp. vii, 349. $6.25.

Here is a book which invites reading: it has excellent print and format, clear-cut organization and unusually attractive and appropriate photographs. The reading of this book, however, which was written by a teacher of child development in home economics education, brings disappointment. From the point of view of this reviewer, there are two reasons for this.

First, although the book is presumably focussed on the guidance of young children, the emphasis is much more on *what* children do than on *how* children learn, and on *what* teachers should do rather than on the *why* or the *how* of their doing. Even the chapter headings, except for Chapter 3, do not refer to guidance. This does not mean that there are no specific suggestions for guidance because there are many, *e.g.* see p. 114 on dealing with a child's aggression, but by and large there is too much reliance on general admonitions that the teacher should "try to determine the feelings of the child" (p. 144), or "look for underlying causes" (p. 131), without the background necessary to help the student learn how such difficult tasks may be done.

Secondly, the developmental and psychological material is presented on a superficial level. There is no documentation of statements and little or no reference to the theories of specialists in the field. The list of readings at the close of each chapter is chosen from secondary sources, with the exception of a reference to a single research study at the end of a few chapters for "students who plan to do further study in child development." The author states that this book is for beginning students and "is not intended as a detailed and scientific account of human development."

A serious lack in this book, which was written primarily for college students is the failure to introduce the student to the studies of child psychologists and cultural anthropologists. The lack of such stimulating materials leaves the student dependent upon the authority of the author, rather than with increased ability to read and judge for himself, which is the aim of education on the college level. Most beginning students could tackle with zest stronger fare than this book offers.

There are numerous statements in the book which are questionable, *e.g.* "a child's personality will be changing incessantly" (p. 28); "children, in releasing their pent-up anger, may try to hurt inanimate objects. This is proof of the impersonal attitudes that underlie many of their aggressive acts" (p. 68); and "the value of setting regular times when a lull in nursery school activity provides an opportunity for toileting" (p. 156). For the sake of brevity, I shall, however, comment only on her statement that "the major aim of guidance is to help children learn to live happily with other people." I would certainly agree that this is one major part of the purpose of a teacher's guidance, but should we not add to this the aim of stimulating the perceptions of children and helping them in the development of understanding? A short paragraph on p. 126 plus a brief statement on pp. 230-231 is all that the author devotes to a discussion of mental development. It seems hardly

adequate to orient the student to the complexities of the young child's mind as he strives to perceive and understand the cultural and physical world in which he lives. The area of perceiving and thinking is scarcely touched upon in this book.

The author is at her best, it seems to me, when she discusses children's play and the environment which the nursery school provides to stimulate and guide activities through creative materials, stories, music, and experiences in nature and science. The beginning teacher and the beginning mother will find many fruitful suggestions here.

Stanford University LOIS MEEK STOLZ

The Psychology of Deafness. By HELMER R. MYKLEBUST. New York, Grune and Stratton, 1960. Pp. xii, 393. $7.75.

Rarely do we find a book on the sensorially-impaired that is of interest to the general psychologist. Most of the research effort in the field seems to have been expended on the practical problem of the assessment of behavior which is of concern primarily to the professional worker. The present volume on the psychology of deafness, however, does attempt to go beyond the presentation of empirical data to derive implications for general psychological functioning. By frequent reference to comparisons between the normal and the deaf and by partialling out the effects of sex, age, etiology, and duration of hearing, the author strives to reveal the basic ways in which hearing influences perceptual organization. At the same time, whenever feasible, the implications for learning, education, and personality development are also presented. Nowhere is this better illustrated than in the chapter on mental development. Gross quantitative differences (or lack of differences) between normals and those whose hearing is impaired are probed for qualitative differences. The author is careful to point out the difficulties in drawing conclusions about the deaf based on tests standardized on hearing groups.

Unfortunately, the problem of cross-validation is often ignored in the decisions concerning the significance of item analyses. Also simple *t*-tests on the mean differences of subgroups are conducted without first demonstrating the significance of the overall *F-ratio*. Too much space is devoted to the results of psychometric and projective tests, some of doubtful validity. For example, the analyses of the data on drawing tests take up 21 pages in the two chapters on mental and emotional development. The chapter on personality development relies too heavily on the results of the *MMPI*. Perhaps the emphasis in the book on test-methodology only reflects the characteristic preoccupation of investigators in the field. That nature has provided a laboratory for the study of sensory contributions to psychological processes is often ignored. It is for this reason, perhaps, that greater attention should have been given to experimental investigations, particularly in such areas as space perception, orientation, and cognitive processes. The paucity of experimental results, however, does not detract from the value of the book to workers in the field of audiology. The vast amount of research literature, simply and logically presented, covering almost all phases of the behavior of the deaf, will undoubtedly make the present volume a frequent source of reference. The brief introductory section describing the hearing mechanism and summarizing the definitions, types, and extent of deafness, however, could have been left out or more carefully prepared. It is too elementary for "a text in advanced courses" and it often leads to statements of doubtful accuracy (as in referring to the range of frequencies of 20-20,000 cycles to which Man can respond,

or to the five senses, or to the *inability* to stimulate the vestibular mechanism from outside the body). The importance of language in the deaf, both verbal and written, is carefully developed and will certainly prove of great interest not only to the teachers of the deaf but to those psychologists interested in the contributions of hearing to language development. The final chapters on other handicaps and special interests and aptitudes of the deaf, though not germane to the major topic, are welcome additions to the literature and might help stimulate more interest in these areas.

PHILIP WORCHEL
University of Texas

Inner Conflict and Defense. By DANIEL R. MILLER and GUY E. SWANSON, in collaboration with WESLEY ALLINSMITH, ELTON B. MCNEIL, and BEVERLY B. ALLINSMITH, JUSTIN ARONFREED, BETTY J. BEARDSLEE, LEONARD M. LANSKY. New York, Henry Holt and Company, 1960. Pp. x, 452. $6.95.

Several related researches, all a part of a single program, are reported in this volume. That this program was not narrowly focused is indicated by the titles used for the three sections on results: moral standards, mechanisms of defense, and expressive styles. There is a greater unity than these titles might immediately suggest, however; nearly everything in the book could be said to be related to conflict, and the research is all directed at understanding the effects of social structure on conflict resolution in contemporary American society.

The research deals only with one sex. The principal subjects were boys in the seventh to ninth grades. Studies involving sex identity required older subjects, and university students were used. The social structure variables were social class (middle vs. working) and type of social integration (entrepreneurial vs. bureaucratic). Behavioral data were collected directly from the subjects; and information about child rearing, by interviews with the mothers of the boys. The behavioral variables are studied as a function of social structure, and the relation to child training variables is studied by itself and as a function of social structure.

The book which results is indispensable for all scholars working on socialization, on personality development, and on dynamics of adjustment. The findings obtained and the methods used are relevant to many and sometimes discrete lines of inquiry on these general topics. As scholars in these fields are hardly in danger of remaining unaware of so distinguished a contribution, it is perhaps more important to say here also that parts of this book will be relevant to the special interests of psychologists in other fields as well. For example, any student of motivation will be interested in the use of short-term change scores on projective tests (after-arousal compared with before-arousal) as a measure of a long-term personality characteristic, and the substantial evidence of special value in the use of such a measure. The student of personality structure will find a wealth of suggestions. Anyone interested in personality testing will find here new tests, and new light on the meaning of old tests. Even these statements, however, do not adequately indicate the great variety of points of view in behavioral science from which one or another part of this volume alters the present scene.

IRVIN L. CHILD
Yale University

Books Received

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

REVISED EDITIONS

de Gutierrez-Mahoney, C. G., and Carini, Esta, *Neurological and Neurosurgical Nursing*, 3rd ed., St. Louis, C. V. Mosby Co., 1960, pp. 413, $6.50.

Lachman, Sheldon J., *The Foundations of Science*, rev. ed., New York, Vantage Press, 1960, pp. 121, $2.95.

Newton, Kathleen, *Geriatric Nursing*, 3rd ed., St. Louis, C. V. Mosby Co., 1960, pp. 483, $6.50.

Pigors, Paul, and Myers, Charles A., *Personnel Administration. A Point of View and a Method*, 4th ed., McGraw-Hill Book Company, 1961, pp. x, 749, $7.95.

Seidman, Jerome M., *The Adolescent. A Book of Readings*, rev. ed., New York, Henry Holt and Co., 1960, pp. x, 870, $6.50.

Ulett, George A., and Goodrich, W. Wells, *A Synopsis of Contemporary Psychology*, 2nd ed., St. Louis, C. V. Mosby Co., 1960, pp. 309, $6.50.

Williams, Robin, Jr., *American Society. A Sociological Interpretation*, 2nd ed., New York, Alfred A. Knopf, 1960, pp. xv, 575. xiv.

SOFT-COVERED REPRINTS

Behanan, Kovoor T., *Yoga. A Scientific Evaluation*, New York, Dover Publications, 1937, pp. xx, 270, $1.65.

de Ropp, Robert S., *Drugs and the Mind*, New York, Grove Press, 1960, pp. x, 310, $1.95.

Evans-Wentz, W. Y., *The Tibetan Book of the Dead*, New York, Oxford University Press, 1960, pp. xxi, 249, $1.75.

Freud, Sigmund, *Group Psychology and the Analysis of the Ego*, translated by Strachey, James, with an introduction by Franz Alexander, New York, Bantam Books, 1960, pp. xvii, 108, 50¢.

Freud, Sigmund, *Beyond the Pleasure Principle*, translated by Strachey, James, with introduction and notes by Gregory Zilboorg, New York, Bantam Books, 1959, p. 121, 50¢.

Krich, A. M., (ed.), *The Anatomy of Love*, New York, Dell Publishing Co., 1960, pp. 317, 50¢.

Lindzey, Gardner, (ed.), *Assessment of Human Motives*, New York, Grove Press, 1960, pp. 273, $1.95.

Malinowski, Bronislaw, *A Scientific Theory of Culture and Other Essays*, New York, Oxford University Press, 1960, pp. ix, 228, $1.50.

McDougall, William, *An Introduction to Social Psychology*, New York, Barnes and Noble, 1960, pp. xxix, 524, $1.95.

Rashevsky, N., *Mathematical Biophysics. Physico-Mathematical Foundations of Biology*, New York, Dover Publications, 1960, vol. 1, pp. xxvi, 488, $2.50; vol. 2, pp. xii, 462, $2.50.

Rawcliffe, D. H., *Illusions and Delusions of the Supernatural and the Occult*, New York, Dover Publications, Inc., 1960, pp. 551, $2.00.

Reik, Theodor, *The Haunting Melody*, New York, Grove Press, 1960, pp. viii, 376, $2.45.

Reik, Theodor, *The Secret Self. Psychoanalytic-Experiences in Life and Literature*, New York, Grove Press, 1960, pp. 329, $2.45.

Sumner, William Graham, *Folkways*, New York, New American Library of World Literature, Inc. (Mentor book), pp. xvi, 605, 75¢.

Walker, Nigel, *A Short History of Psychotherapy in Theory and Practice*, New York, Noonday Press, 1960, pp. xii, 185, $1.45.

Wolberg, Lewis R., *Hypnoanalysis*, New York, Grove Press, 1960, pp. xviii, 342, $2.25.

Zilboorg, Gregory, *Sigmund Freud*, New York, Grove Press, 1960, pp. 132, $1.45.

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXXIV DECEMBER, 1961 No. 4

SPREAD OF EFFECT AS A FUNCTION OF TIME AND INTRASERIAL SIMILARITY

By LEO POSTMAN, University of California, Berkeley

Thorndike's hypothesis that the spread of effect reflects the automatic action of reward has been seriously questioned. In his original report Thorndike had described a symmetrical gradient of repetition around the rewarded position.[1] Later investigations have rarely found a significant fore-gradient. An after-gradient has usually been found but has been widely attributed to biases in the sequence of S's-responses.[2] According to that interpretation, the repetition of the rewarded response initiates a non-random sequence of further responses. Since the sequential dependencies between successive responses are less than perfect, a gradient of repetition following the rewarded response is generated. There is substantial independent evidence for the operation of such sequential biases.[3]

Two experimental findings appear to support the view that, in a 'Thorndikian' situation (*e.g.* when *S* guesses numbers in response to words), sequential biases account for the spread of effect. The first critical finding comes from experiments in which the positions of the stimuli were shifted between training and test, with only the serial positions of the rewarded pairs held constant. In spite of the shifts

* Received for publication May 12, 1960. This study was supported by a grant from the National Institute of Mental Health. The writer is grateful to Miss Polly Davis for her assistance.

[1] E. L. Thorndike, An experimental study of rewards, *Teach. Coll. Contr. Educ.*, No. 580, 1933, 7-56.

[2] W. O. Jenkins and F. D. Sheffield, Rehearsal and guessing as sources of the 'spread of effect,' *J. exp. Psychol.*, 36, 1946, 316-330; W. O. Jenkins and L. M. Cunningham, The guessing-sequence hypothesis, the 'spread of effect,' and number-guessing habits, *ibid.*, 39, 1949, 158-168; M. H. Smith, Spread of effect is the spurious result of non-random response tendencies, *ibid.*, 39, 1949, 355-368.

[3] Jenkins and Cunningham, *op. cit.*, 164-167; Sheffield, 'Spread of effect' without reward or learning, *J. exp. Psychol.*, 39, 1949, 375-379; C. A. Fagan and A. J. North, A verification of the guessing sequence hypothesis about spread of effect, *ibid.*, 41, 1951, 349-351; M. H. Marx and Ben Bernstein, Generalization and reinforcement among responses to synonyms, *J. gen. Psychol.*, 52, 1955, 49-64.

in the positions of the stimuli, there was a gradient of repetition for the *step-positions* following the rewarded response. No clear after-gradient was found when repetitions were determined for specific stimulus-response pairs which had occurred at varying distances from the reward during training.[4] Apart from some methodological questions which can be raised about these studies,[5] it is doubtful whether the procedure of shifting the positions of the stimuli permits an adequate analysis of the 'Thorndike effect.' The hypothesis at issue is that the effects of reward spread to neighboring, punished stimulus-response associations. For purposes of analysis, the term, *stimulus,* should refer not only to the item presented to *S* but also to contextual features, notably the position of the item in the series. Thus, when the positions of the stimuli are shifted, each item on the test appears in a novel serial context. What is tested, then, is not repetition of a response to a recurrent stimulus-complex but rather generalization to a substantially changed stimulus-complex. Such a procedure may well be insensitive to spread of effect as originally described by Thorndike.

A second finding apparently favorable to the 'guessing-sequence' hypothesis is the dependence of the after-gradient on repetition of the rewarded response.[6] The sequence of non-random guesses is assumed to be anchored to the rewarded response and should not occur unless that response is repeated. Demonstrations of this relationship have not been conclusive. When the responses of all *S*s are pooled and the gradients following repetition and non-repetition of the rewarded response are compared, it is probable that all *S*s do not contribute equally to each of the two gradients.[7] In a random sample of *S*s there usually is considerable variation in speed of learning as measured by repetitions of the rewarded responses. Thus, fast learners necessarily contribute more than do slow learners to the gradient following repetition of the rewarded response; the converse is true for the gradient following non-repetition of the rewarded response. There are at least two ways in which the selection of *S*s may influence the difference between the two gradients: (a) if speed of learning reflects effectiveness of the rewards given during training, there may be correlated differences in the amount of spread; (b) if speed of learning is inversely related to the amount of intraserial interference, competition from misplaced responses may reduce the gradient more for slow than for fast *S*s. It is necessary, therefore, to examine the dependence of the after-gradient on repetition of the rewarded response by methods which avoid the selection of *S*s.

The empirical findings usually cited in support of the guessing-sequence hypothesis thus appear to be equivocal. This conclusion must be considered in conjunction with Marx's demonstration of a significant spread of effect based on a comparison between gradients following repeated rewarded and repeated non-rewarded re-

[4] G. A. Zirkle, Success and failure in serial learning: I. The Thorndike effect, *J. exp. Psychol.*, 36, 1946, 230-236; Jenkins and Cunningham, *op. cit.*, 158-164.

[5] These questions are discussed in Marx, Spread of effect: a critical review, *Genet. Psychol. Monogr.*, 53, 1956, 135-137.

[6] Jenkins and Sheffield, *op. cit.*, 322; Jenkins and Cunningham, *op. cit.*, 160-162; Marx, Gradients of error-reinforcement in a serial perceptual-motor task, *Psychol. Monogr.*, 71, 1957, No. 437, 6 f.

[7] Such an analysis was made by Jenkins and Sheffield, *loc. cit.*, and Jenkins and Cunningham, *loc. cit.* An alternative procedure is that used by Marx, 1957, *loc. cit.*, in which *S*s are divided into sub-groups according to the amount of variation in the rewarded response.

sponses, *i.e.* with the effects of guessing-sequences held constant.[8] Such evidence no longer permits the reduction of the spread of effect to the operation of sequential response-biases.

The experiment reported in this paper was designed to re-assess the contribution of sequential biases to the spread of effect in a classical 'Thorndikian' situation. The primary purpose of the study is to make an unbiased test of the alleged dependence of the after-gradient upon repetition of the rewarded response. For purposes of manipulating the frequency of repetition of the rewarded response without recourse to selection of *S*s, the spread of effect was measured immediately after the end of training and after an interval of time. The question was whether the forgetting of correct responses in the rewarded positions would be accompanied by a proportional reduction in the after-gradient. In view of the probable interaction between the effects of reward and the conditions of generalization within the series, the experiment was replicated with materials of high and low intraserial similarity. The finding of Duncan that gradients of repetition are influenced by intra-list similarity indicates that the conditions of stimulus-generalization should be considered in any general analysis of the determinants of the spread of effect.[9]

METHOD

Original learning. A series of 21 adjectives was presented to *S* who was instructed to guess a number between 2 and 9 in response to each word. The stimulus-words were those used by Duncan. There were (a) a high-similarity list of synonymous adjectives, *i.e. best, perfect, supreme,* etc.; and (b) a low-similarity list of unrelated adjectives.[10] The responses in the 5th, 11th, and 17th positions were always called right, and all other responses were called wrong. Four different random orders were used equally often with each list. The following restrictions were used in the construction of the different orders: (a) no word occupied a rewarded position in more than one series; (b) no word occurred in the same position more than once; (c) the same four words were always used in the two initial and two terminal positions although their specific positions varied from series to series. Since five punished responses intervened between successive rewards, measurement of the fore-gradient and after-gradient was limited to two positions each.

---

[8] Marx, *op. cit.,* 1957; M. H. Marx and F. E. Goodson, Further gradients of error reinforcement following repeated reinforced responses, *J. exp. Psychol.,* 51, 1956, 421-428.

[9] C. P. Duncan, Stimulus-generalization and spread of effect, this JOURNAL, 64, 1951, 585-590.

[10] Duncan, *op. cit.,* 586. One word was added to each of Duncan's lists of 20 words, so that there were two 'buffer' items both at the beginning and the end of the series. The responses were restricted to the range 2-9 since the numbers 1 and 10 appeared to have pre-experimental associations with the items in the high-similarity list.

The stimulus-words were presented on a memory-drum at a 4-sec. rate. Each word was exposed for 2 sec., and *S* called out his response during that period. During the remaining 2 sec. of the cycle, the word was covered by a shutter, and *E* made his announcement of 'right' or 'wrong.' Prior to presentation of the stimulus-list, *S* practiced the procedure by calling out letters of the alphabet in response to a series of typed symbols.

The instructions informed *S* that the list would be presented more than once, and that he was to repeat responses which had been called right and to avoid repetition of responses which had been called wrong. The instructions also stated that more than one number might be correct in response to a given stimulus-word.

Tests of retention. Four test-trials were given either 30 sec. or 20 min. after

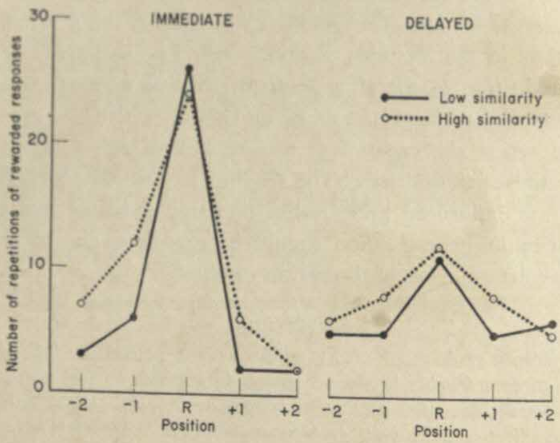


FIG. 1. FREQUENCY OF REPETITIONS OF REWARDED RESPONSES AS A FUNCTION OF DISTANCE FROM CORRECT POSITION

the end of the original training-trial. The procedure on the test-trials was exactly the same as on the training-trial. The interval between successive test-trials was 4 sec. For the groups given a delayed test, the 20-min. interval was filled with practice on a pursuit-meter.

Subjects. With two conditions of intra-list similarity and two retention-intervals, the experimental design comprised 4 groups of 16 *Ss* each. Assignment to experimental groups was in blocks of 4 *Ss*, one to each of the combinations of intra-list similarity and retention-interval. The order of conditions within each successive block was determined by a table of random numbers as was the assignment of the different serial orders of each list.

RESULTS

(1) Frequency of repetitions. All measures of repetition are based on the results of the first test-trial. The numbers of rewarded responses re-

peated are shown in Fig. 1 (Position R). The frequency of repetitions is approximately the same for the two conditions of intra-list similarity. There is a considerable and significant amount of forgetting of rewarded responses during the 20-min. interval ($F = 13.97$, $df. = 1$ and 60, $P < 0.01$).[11] The failure to find a difference in retention as a function of stimulus-similarity agrees with the results obtained in studies using the conventional method of paired-associate learning.[12] For punished responses (Fig. 2) there is no reliable drop in the number of repetitions after the 20-min. interval ($F = 1.16$). The apparent difference in the

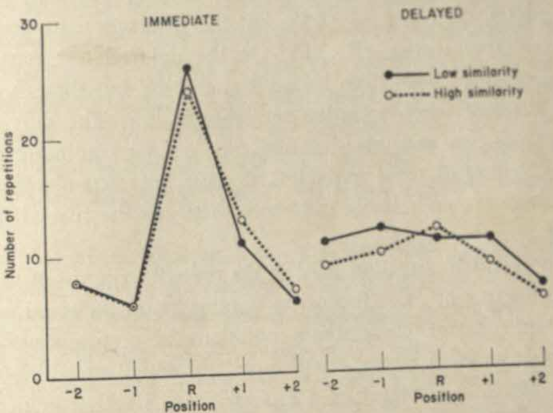


FIG. 2. FREQUENCY OF REPETITIONS OF RESPONSES AS A FUNCTION OF DISTANCE FROM REWARDED POSITION

rate of forgetting for rewarded and punished responses is probably largely a matter of regression. Whatever the reasons for the reduction of the difference between Positions R and +1, the implications for the recurrence of non-random sequences of responses remain the same.

(2) Intra-list errors. Fig. 1 shows, in addition to the number of correct repetitions, the frequency with which the rewarded response was given in the two positions preceding and following the correct one. By analogy to the intra-list errors observed in serial anticipation, these misplaced rewarded responses may be designated as anticipatory and perseverative er-

[11] Prior to analysis of variance, the scores were subjected to a Freeman-Tukey square-root transformation. This transformation was used for all analyses of measures obtained on the first test-trial.

[12] B. J. Underwood, Studies of distributed practice: IX. Learning and retention of paired adjectives as a function of intralist similarity, *J. exp. Psychol.*, 45, 1953, 143-149.

rors, respectively. Since we are concerned with differences as a function of position and intra-list similarity, the frequency of such errors to be expected by chance need not be considered.

On the immediate test, misplaced repetitions of the rewarded response occur more frequently for the high-similarity than for the low-similarity list. Anticipatory errors are more numerous than perseverative errors. When intra-list similarity is high, both types of error describe a gradient; when intra-list similarity is low, there is a gradient only for anticipatory errors. The difference between the two lists is significant ($F = 4.98$, $df. = 1$ and 30, $0.02 < P < 0.05$), as is the variation among positions ($F = 3.63$, $df. = 3$ and 90, $0.02 < P < 0.05$), but the interaction of position with intra-list similarity is not ($F < 1$). On the delayed test, a symmetrical gradient is obtained with high intra-list similarity, but there is no indication of a gradient when intra-list similarity is low. The differences as a function of position and intra-list similarity are not significant on the delayed test. Neither of these variables, however, interacts significantly with time, *i.e.* there are no reliable differences between the immediate and delayed tests.

The 'Thorndikian' situation used in the present experiment may be considered an instance of paired-associate learning, with the serial positions of the pairs held constant. Under these conditions the occurrence of misplaced correct responses is to be expected, especially when intra-list similarity is high. The fact that anticipatory errors are more likely than perseverative errors is in keeping with the results obtained in serial learning.[13] It is important to recognize the implications of these findings for the measurement of the spread of effect. Misplaced rewarded responses may compete effectively with punished responses in the positions surrounding the rewarded pair. To the extent that the misplaced repetitions of the rewarded response describe a gradient they may serve to mask the spread of effect. The results shown in Fig. 1 suggest that the competition between misplaced rewarded responses and punished responses (a) is a function of intra-list similarity, (b) is greater in the positions preceding, than in the positions following, the reward, and (c) may be more effective in an immediate than in a delayed test. All the conditions favorable to such competition should be detrimental to the detection of the spread of effect.

(3) Spread of effect. Evidence concerning the spread of effect will be considered under three headings: (a) after-gradients of repetition on the

[13] See, for example, B. R. Bugelski, A remote association explanation of the relative difficulty of learning nonsense syllables in a serial list, *J. exp. Psychol.*, 40, 1950, 336-348.

first test-trial, (b) fore-gradients of repetition on the first test-trial, and (c) variability of responses during the five trials.

(a) After-gradients of repetition. As Fig. 2 shows, there is a clear after-gradient of repetition of errors on the immediate test under both conditions of similarity. For the difference between Positions +1 and +2, $F = 4.62$ ($df. = 1$ and 30, $0.02 < P < 0.05$). The results for the two lists are very much alike, *i.e.* the pattern of repetitions is not influenced by intra-list similarity. On the delayed test, repetitions of the rewarded responses decline drastically, so that there is little or no difference between Positions R and +1. Nevertheless, an after-gradient still appears to be present. Considered by itself, the after-gradient on the delayed test is not reliable ($F = 1.83$).[14] Neither position nor similarity, however, interacts significantly with time ($F < 1$ in both cases). Thus, the statistical tests point to an after-gradient which is not reduced significantly by the delay between training and test. Additional analysis of *S*s' response-sequences will serve to strengthen the conclusion that an after-gradient persists on the delayed test.

The results of the delayed test show that the after-gradient is not eliminated even when rewarded responses are repeated no more frequently than are errors in Position +1. This finding, obtained with unselected samples of *S*s, indicates that the after-gradient is not contingent on the repetition of the rewarded response. The same conclusion is supported by an examination of the specific sequences of responses given by individual *S*s both on the immediate and the delayed test. For the three positions, R, +1, and +2, there are eight possible sequences of repetitions and non-repetitions of responses. Table I shows for both tests the frequencies with which each of these sequences occurred, and the numbers of *S*s contributing to each of these frequencies. The two sequences which are critical for purposes of the present analysis are YYN and NYN. Both sequences contribute to an after-gradient; the former is anchored to a repeated rewarded response whereas the latter is not. On the immediate test, the two sequences occur with about equal frequency. On the delayed test, NYN is a more frequent sequence than is YYN. The results for the two conditions of intra-list similarity are in close agreement. The findings are

[14] Two supplementary analyses show that the pattern of repetitions on the delayed test was not a chance one. First, the rewarded response was repeated significantly more often in Position R than in the adjacent positions (Fig. 1). Secondly, the repetitions were significantly greater in Positions +1 and +2 than in serial positions 8, 14, and 20 which were three steps removed from a reward and not included in the analysis of the gradients.

the same whether we consider the total frequencies of the sequences or the numbers of *S*s contributing them.

The shift on the delayed test in favor of class NYN undoubtedly reflects the change in opportunities for the two kinds of sequences entailed by the drop in repetitions of rewarded responses. In fact, such a shift is implied by the persistence of the gradient in spite of a large drop in repetitions of rewarded responses. The important point is that sequences of the class NYN contribute equally with those of class YYN to the gradient on the immediate test and are primarily responsible for the gradient on the delayed test. Thus, analysis of the individual response-

TABLE I

DISTRIBUTION OF SEQUENCES OF REPETITIONS AND NON-REPETITIONS IN POSITIONS R, +1 AND +2

(Y = Yes = Repetition; N = No = Non-repetition; *f* = Frequency of sequence; *S*s = Number of *S*s contributing each frequency.)

	Immediate test				Delayed test			
Sequence	high similarity		low similarity		high similarity		low similarity	
R+1+2	*f*	*S*s	*f*	*S*s	*f*	*S*s	*f*	*S*s
Y Y Y	0	0	1	1	0	0	1	1
Y Y N	*6*	*5*	*4*	*4*	*1*	*1*	*2*	*2*
Y N Y	2	2	2	2	0	0	2	2
Y N N	16	9	19	13	11	9	6	6
N Y Y	1	1	0	0	1	1	2	2
N Y N	*5*	*5*	*6*	*6*	*7*	*7*	*6*	*6*
N N Y	3	3	3	3	5	5	2	2
N N N	15	10	13	9	23	13	27	15

sequences yields no evidence for the dependence of the after-gradient on repetition of the rewarded response.[15] It should also be noted that Table I provides no indication that repetitions in Position +2 are contingent on repetitions in Position +1.

(b) Fore-gradients of repetition. As Fig. 2 shows, there is no fore-gradient of repetition of errors on the immediate test. In fact, the number of repetitions is somewhat higher in Position −2 than in Position −1. There are some slight indications of a fore-gradient on the delayed test, but the difference between positions does not approach statistical significance. The interaction of time with position is also not significant.

The absence of a fore-gradient on the immediate test should be considered in conjunction with the gradient of anticipatory errors shown in Fig. 1. To the extent that anticipatory repetitions of the rewarded response

[15] Duncan (*op. cit.*, 589) also failed to find such a dependent relationship.

compete with the punished responses in Positions −1 and −2, a fore-gradient of repetition of errors may be masked. The fact that there is a tendency toward a reverse gradient agrees with this interpretation since the competition from anticipatory rewarded responses should be substantially greater in Position −1 than in Position −2. On the delayed test, the gradient of anticipatory rewarded responses is reduced, and the reversal in the error-gradient is no longer present. While the data are consistent with this analysis, it is not possible to assert with any confidence that a fore-gradient of repetition of errors was, in fact, being masked by anticipatory repetitions of rewarded responses. It is useful to recognize, however, that anticipatory repetitions of the rewarded response complicate the measurement of the fore-gradient.

(c) Gradients of variability. If there is a gradient of repetition of errors, the number of *different* responses given during a series of trials should vary inversely with the distance from the rewarded position. The number of different responses, which is an index of response-variability, has been shown to provide a sensitive measure of the spread of effect. Thus, in Duncan's study this index yielded a significant fore-gradient under conditions of high intra-list similarity whereas the conventional measure of repetitions did not.[16] As a measure of spread, the index of response-variability has the advantage of being based on a series of trials, *i.e.* it is likely to be more reliable than the number of repetitions on a single trial. In addition, a measure of response-variability will be sensitive to the gradient of 'discontinuous' repetitions of errors described by Marx and Bunch.[17] These investigators reported that the tendency to repeat errors made more than one trial earlier was inversely related to distance from the rewarded position. It must be recognized, however, that the cumulative effects of associative interference may serve to reduce the sensitivity of the index.

Table II shows the mean numbers of different responses per step-position given during five trials (one original training-trial and four test-trials). In Position R, there is a significant increase in the number of different responses as a function of time ($F = 9.17$, $df. = 1$ and 60, $P < 0.01$). This increase in variability reflects the drop in repetitions of the rewarded responses after the 20-min. interval. The effects of the delay do not interact with intraserial similarity ($F < 1$).

As measured by the index of response-variability, there is an after-gradient which is equally steep on the delayed and immediate tests. An

[16] Duncan, *op. cit.*, 587 f.

[17] M. H. Marx and M. E. Bunch, New gradients of error reinforcement in multiple-choice human learning, *J. exp. Psychol.*, 41, 1951, 93-104.

over-all analysis of variance shows the different between Positions +1 and +2 to be highly significant ($F = 8.30$, $df. = 1$ and 60, $P < 0.01$). This difference interacts neither with time nor with intra-list similarity ($F < 1$ in both cases). This analysis provides strong support for the conclusion reached earlier that the after-gradient persists on the delayed test, in spite of the substantial reduction in the repetition of rewarded responses. The fact that the after-gradient of response-variability is as steep on the delayed test as it is on the immediate test is worthy of emphasis.

The index of variability fails to produce clear evidence for a fore-gradient. The differences between Positions −1 and −2 are negligible with one exception, viz. the delayed test of the group learning the high-similarity list. Analysis of variance shows that the differences on the immediate test do not approach significance. There are no significant inter-

TABLE II

Mean Number of Different Responses per Step-Position Given during Five Trials

Step-position	Immediate test				Delayed test			
	high similarity		low similarity		high similarity		low similarity	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
−2	3.79	0.32	3.73	0.56	3.87	0.54	3.83	0.35
−1	3.71	0.45	3.79	0.47	3.58	0.47	3.81	0.41
R	2.33	0.73	2.16	0.62	2.77	0.55	2.71	0.62
+1	3.56	0.69	3.56	0.45	3.60	0.42	3.54	0.40
+2	3.73	0.44	3.85	0.48	3.81	0.45	3.85	0.52

actions when the two tests are compared. Considered by itself, the fore-gradient on the delayed test for the high-similarity list is significant ($t = 2.48$, $df. = 15$, $0.02 < P < 0.05$). In view of the results of the analysis of variance, this single test must be interpreted with caution. It is important to note, however, that the largest difference between Positions −1 and −2 is obtained under conditions which should be most favorable to the detection of a fore-gradient, *i.e.* (a) high intra-list similarity which is unfavorable to the discrimination of simulus-response sequences, and (b) a delayed test in which there is relatively little competition from misplaced repetitions of the rewarded response.

Discussion

Before discussing the general implications of the experiment, we shall briefly compare our findings with those of Duncan.[18] Only the results of the immediate tests are relevant to this comparison since Duncan's data

[18] Duncan, *op. cit.*, 585-590.

were obtained in eight consecutive trials. It should be noted, however, that even with respect to the immediate tests, there are several differences between Duncan's procedure and ours: (a) In the present experiment 3/21 responses were called right and 4/20 in Duncan's. (b) In reporting the frequency of repetitions, Duncan combined the results of seven test trials, counting as a repetition the occurrence of the same response on a pair of successive trials. We measured the gradient of repetition only on the first test-trial. (c) The rate of presentation was slower in the present study (4 sec.) than in Duncan's (3 sec.).

The after-gradients in the present experiment are steeper and more clearly significant than those reported by Duncan. We find only a negligible difference in after-gradients as a function of intra-list similarity whereas in the earlier study there was an indication of a gradient only when similarity was high. In full agreement with Duncan, we find that the after-gradient is not dependent on repetition of the rewarded response. The two studies also agree in not finding a significant fore-gradient of repetitions. The variability of responses proved to be a sensitive measure of the spread of effect in both experiments. In Duncan's study, the gradient of variability tended to be symmetrical. On the immediate test in the present experiment, only an after-gradient was found; a fore-gradient emerged on the delayed test. In Duncan's experiment, the variability of responses was sampled over a larger number of trials (eight) than in ours (five). Thus, the former provided a better opportunity to detect a fore-gradient of variability on an immediate test. In view of the differences in procedure and measures, the degree of agreement between the two experiments may be considered satisfactory.

The single most important conclusion which can be drawn from the present experiment is a negative one. Our results show clearly that the after-gradient of repetition is not contingent upon repetition of the rewarded response. The absence of a contingent relationship is brought out by (a) analysis of the individual sequences of responses contributing to the after-gradients, and (b) the persistence of the gradients on the delayed test in spite of a sharp decline in the frequency of repeated rewarded responses. The gradients based on frequency of repetitions on the first test-trial and on inter-trial variability of responses are consistent. The persistence of the after-gradient during the temporal interval is brought out especially clearly by the index of variability. Thus, a critical implication of the guessing-sequence hypothesis is not supported by the facts. While sequential biases may contribute to the spread of effect, they cannot account for it. This conclusion is in agreement with that of Marx who found that

reward significantly enhances the after-gradient when guessing-sequences are controlled.[19]

Our analysis shows that there are gradients of repetition not only for errors but also for rewarded responses. Competition from misplaced rewarded responses may serve to reduce the repetition of errors. To the extent that the two gradients are parallel each may be reduced. It is possible that high intra-list similarity did not produce a steeper after-gradient of errors than low intra-list similarity because the competing after-gradient of misplaced rewarded responses was higher and steeper under the former condition. This analysis also suggests that fore-gradients of repetition of errors may be more difficult to detect than after-gradients because they are more subject to competition from misplaced rewarded responses.

The results of the experiment fail to support not only the guessing-sequence hypothesis but also the hypothesis of serial response-response reinforcement which assumes dependence of the after-gradient on repetition of the rewarded response.[20] The data do not offer direct support to an alternative interpretation but are consistent with Thorndike's hypothesis of automatic action of reward. The spread of effect cannot be dismissed as an artifact resulting from sequential response-biases. The mechanism responsible for the gradient of repetition of errors continues to pose a theoretical problem.

Summary

This study has reëxamined the conditions determining the occurrence of the spread of effect in a 'Thorndikian' situation. According to the guessing-sequence hypothesis, the after-gradient is dependent on repetition of the rewarded response to which a non-random sequence of responses is anchored. Much of the evidence in accord with this prediction has been obtained in situations differing from those used by Thorndike with respect to the conditions of training and reinforcement. In addition, tests of the dependence of the gradient on repetition of the rewarded response may have been biased by selection of *S*s, since repetition of both rewarded responses and errors may be a function of learning ability.

To manipulate frequency of repetition of rewarded responses without selection of *S*s, the spread of effect was measured for different groups immediately after the end of training and after an interval of 20 min. A typical 'Thorndikian' situation was used, in which *S*s were differentially reinforced for guessing numbers to words. For purposes of assessing the

[19] Marx, *op. cit.*, 1957, 6 f.; Marx and Goodson, *op. cit.*, 424-425.
[20] Marx, *op. cit.*, 1956, 149 f.

interaction between the effects of reward and the conditions of generalization, the experiment was replicated with materials of high and low intraserial similarity.

A significant after-gradient was found on the immediate test. This gradient, as measured by frequency of repetitions and by variability of responses, persisted on the delayed test in spite of a drastic drop in the repetition of rewarded responses. Analysis of individual sequences of responses also failed to provide evidence for dependence of the after-gradient on repetition of the rewarded response. There were no fore-gradient on the immediate test, and only inconclusive evidence for such gradients on the delayed test. The results did not vary significantly as a function of intra-list similarity. Gradients of misplaced rewarded responses may, however, mask the effects of intra-list similarity on the repetition of errors. There is also some evidence that competition from misplaced rewarded responses reduces the repetition of errors in the positions preceding reward.

The results fail to support the guessing-sequence hypothesis and leave open the possibility that the spread of effect reflects the automatic action of rewards.

ILLUSORY CHANGES IN REPEATED WORDS: DIFFERENCES BETWEEN YOUNG ADULTS AND THE AGED

By RICHARD M. WARREN, National Institute of Mental Health

If one listens to a recording of a word or phrase pronounced distinctly and repeated again and again, illusory changes occur. Warren and Gregory reported that the listener to this repetitive, unchanging stimulus seems to hear the voice shift suddenly to other words and phrases at frequent intervals.[1] These changes may involve considerable phonetic distortion.

A more detailed investigation of this 'verbal transformation' was reported by Warren (*Ss* called out what they heard initially, and then called out each time they thought they heard a change in what the voice was saying during the 3-min. test-period).[2] Verbal transformations were found for all repeating stimulus-words employed (which varied in complexity from "see," with only two phonemes, up to short sentences, such as "Our side is right"). Responses were scored both in terms of numbers of transitions and numbers of different forms reported.[3] This scoring allowed a quantitative comparison of responses obtained under a variety of experimental conditions. The data obtained were limited to a group of young *Ss* —British Naval Ratings in their late 'teens or early twenties.

The present study was undertaken to determine if age-differences exist in susceptibility to the verbal transformation effect. Such information might help in understanding the mechanisms underlying this illusion and also afford some insight into possible age-differences in perception.

METHOD

Apparatus. A recording was made of *E*'s voice repeating the stimulus on a Magnecorder Tape Recorder P16-BAH set for high speed (15 in./sec.). The recording was made in an acoustically 'dead' room, and, to minimize reverberation further, a low gain was used with the microphone close to the mouth.

A length of tape containing a single statement of the stimulus was cut and then

* Received for prior publication May 26, 1961. I am indebted to Dr. Arnold D. Krugman for help in obtaining some of the aged *Ss* employed in this study.

[1] R. M. Warren and R. L. Gregory, An auditory analogue of the visual reversible figure, this JOURNAL, 71, 1958, 612f.

[2] R. M. Warren, Illusory changes of distinct speech upon repetition—the verbal transformation effect, *Brit. J. Psychol.*, 52, 1961, 249-258.

[3] The number of transitions reported is not related directly to the number of different forms reported; as long as at least two forms are heard any number of transitions may be experienced.

spliced into a loop. Since a loop is not durable enough for long experimental use, it was played back on the Magnecorder by means of accessory equipment constructed for the purpose. A plate containing a playback-head with the drive-capstan immediately adjacent was attached to the recorder. A second plate containing a roller grooved to fit the tape could be moved to keep the loop taut and running over the playback-head at constant tension. The distance from the playback-head to the roller could be adjusted to accommodate loops constructed from strips of tape from 5–30 in. long.

The output from the playback of the loops was recorded on a reel of tape by a Tandberg Model 5 recorder. The output from a white-noise generator then was recorded on a second channel. These reels of tape with the voice on one channel and noise on another served as stimuli and were played back on the Tandberg. *S* listened through matched Permoflux PDR8 headphones fitted with large foam-rubber muffs to minimize external sounds. *E* could deliver the recorded voice to both phones, or the voice to one phone and white noise to the other. When one ear was receiving the voice and the other noise, the two stimuli could be switched to the opposite phones by *E*, instantly reversing the roles of the two ears.

TABLE I

MEAN NUMBER OF VERBAL TRANSFORMATIONS DURING 3 MIN.

Repeating stimulus	Repetitions	Transitions		different forms	
		62–86 yr.	18–25 yr.	62–86 yr.	18–25 yr.
1. Tress	360	4.3	22.8	2.3	5.5
2. See	390	4.7	35.8	2.2	5.4
3. Flime (nonsense)	330	6.6	32.4	3.5	7.9
4. Police	338	5.8	33.9	2.2	5.4
5. Trice	320	6.5	29.4	3.0	6.5
6. Our cider's ripe	116	1.8	11.6	1.7	2.7

Subjects. The group of young *Ss* consisted of 16 men and 4 women, all Junior College students from 18–25 yr. old. The group of aged *Ss* consisted of 16 men and 4 women from 62–86 yr. old. Most of the aged *Ss* were retired professional men, who had been employed in the United States Civil Service, and their wives.

Instructions. The following typewritten instructions were read by all *Ss*:

Tell me what the voice seems to be saying as soon as you can. After you call out the word or words, keep on listening carefully. If you think you hear a change to different words, call out what they are at once. Don't worry about whether you are right or not—what I want to know is what the words seem like to you under these special conditions.

Don't forget to call out immediately *every time* you hear a change, whether to something new, or whether the change is back to some words you called out before.

Procedure. After the *Ss* read the instructions, the headphones were put on and adjusted to fit properly. The six stimulus-words—five meaningful and one nonsensical—employed are listed in Table I in the order presented. Half the *Ss* (8 men and 2 women in each group) heard the stimulus-series under binaural conditions first. The other 10 *Ss* in each age-group ran through the series monaurally first, white noise being delivered to the ear not receiving the voice. For monaural listening, after 90 sec. the two sound channels were interchanged, each ear receiving

the voice for half the 3-min. test and noise for the other half of the test. The voice was presented at a peak intensity-level of 80 db. and the noise at an intensity-level of 83 db. (Measurements were made with a General Radio Sound Level Meter, Type 1551A, using flat weighting and fast response.) In the binaural series, each ear received the voice at 80 db. for the entire 3-min. test. The stimuli for each presentation increased from 0 to the test-level in 1 sec.

The *S* called out what he was hearing as soon as he could, and then called each change in what the voice seemed to be saying as soon as it occurred. *E* recorded each response on a sheet of paper divided into separate blocks for each 2.5 sec. of the 3-min. test. After the stimulus was shut off, *S* was given a sheet of paper

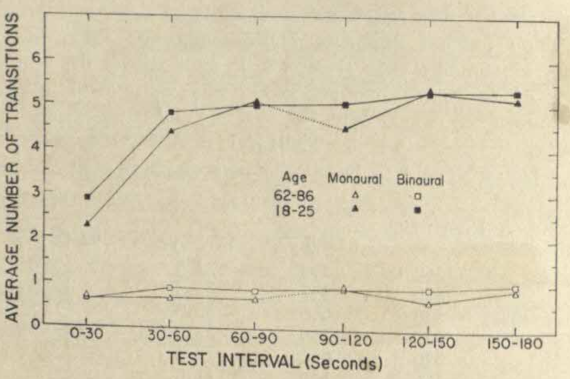


FIG. 1. AVERAGE NUMBER OF ILLUSORY TRANSITIONS REPORTED DURING SUCCESSIVE 30-SEC. INTERVALS

on which he wrote down all the words he could remember having heard. This list was checked against the record transcribed by *E* to ensure that the responses were being interpreted correctly.

It took about 45 min. to complete the first series presented to *S*. There was a rest-interval of 45 min. in which the *S* engaged in activities of his choice, followed by presentation of the second series—each *S* receiving one binaural and one monaural series.

RESULTS

Responses to the individual stimulus-tapes. Table I shows the stimuli used in order of presentation, and the number of times each was repeated during the 3-min. test-period. The average number of transitions and different forms reported by each age-group are summarized. (Monaural and binaural data are combined since, as shown below, both methods of listening produce about the same numbers of responses.) It can be seen from Table I that the younger group reported on the average about five times as many changes as the older group. The younger group also re-

ported more different forms, more than twice as many with most stimuli. While not indicated by Table I, there was in addition a rather marked difference in the way changes occurred for the two groups. Younger *Ss* typically heard abrupt changes from one repetition to the next. With older *Ss*, changes may take place gradually during several repetitions.

Monaural vs. binaural listening. The number of transitions reported with monaural listening tends to be slightly less than with binaural listening for both age-groups, but the difference is not significant for either

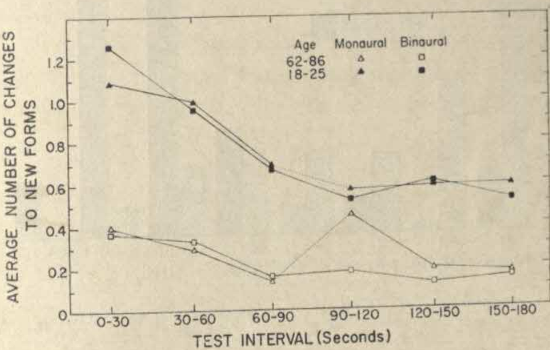


FIG. 2. AVERAGE NUMBER OF CHANGES TO A PREVIOUSLY UNREPORTED FORM DURING SUCCESSIVE 30-SEC. INTERVALS

group. The frequency of monaural and binaural transition (data averaged for all 6 stimuli) during each of the successive 30-sec. intervals is plotted in Fig. 1.

The average number of changes to new forms during each 30-sec. interval is shown in Fig. 2. It can be seen that the numbers of new forms for binaural and monaural listening are equivalent except for the marked increase immediately following the monaural shift in ears (at 90 sec.) for the older group.

Changes in rate and variety of verbal transformations with stimulation-time. It can be seen from Fig. 1 that the frequency of transition stayed at a constant level for the aged during the 3 min. test. The frequency was much higher for the young group, and, except for the first 30-sec. interval, was approximately constant throughout the entire test. One reason for this low score in the first 30 sec. is that the first response was not counted as a transition (as were all subsequent responses), and *S* of necessity spends some time hearing this initial organization of speech sounds before experiencing a transition.

As shown in Fig. 2, the frequencies with which new forms were called

out decreased rapidly at first and then leveled off to a constant rate by 90 sec. (except for a transient increase following the shift in ears during monaural stimulation of the aged group).

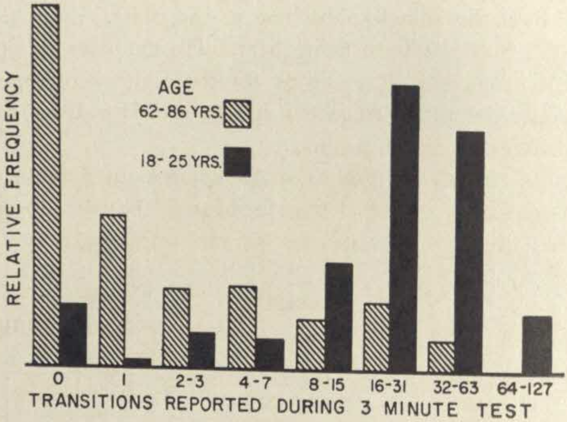


FIG. 3. DISTRIBUTION OF SCORES FOR TRANSITIONS

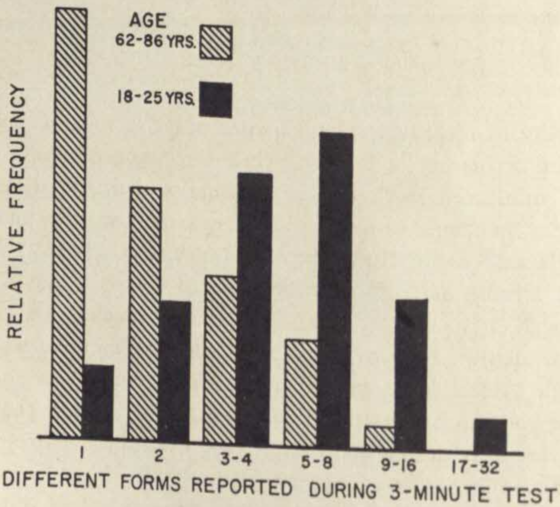


FIG. 4. DISTRIBUTION OF SCORES FOR NUMBERS OF DIFFERENT FORMS

Distribution of scores. The distribution of scores for transitions heard during the 3-min. tests is shown in Fig. 3. For the older group, almost half the tests involved no changes at all—*Ss* hearing the same word

throughout the 3 min. It can be seen that the behavior of the younger group was quite different, the frequency of responses rising to a maximum for the interval representing 16–31 transitions in 3 min.

Fig. 4 shows the distribution of scores for numbers of different forms. For almost half the stimuli presented to the older group, no changes were reported, only one form being heard. For the older group, the frequency of occurrence of scores drops for the higher score-intervals, but for the younger group, the frequency rises to a maximum for the interval from 5–8 different forms in 3 min.

Accuracy of stimulus-identification. In addition to differences in frequency and in variety of verbal transformations, the two age-groups differed in their ability to identify the stimulus-words correctly. The older

TABLE II

FREQUENCY OF USAGE AND CORRECT IDENTIFICATION OF STIMULUS-WORDS

Word	Frequency of usage per million words	Percentage: correct reports, 62–86 yr.	correct reports, 18–25 yr.	time word heard correctly, 62–86 yr.	time word heard correctly, 18–25 yr.
See	>1000	100	100	84	46
Police	70	100	100	74	45
Tress	2	90	95	77	50
Trice	1	75	100	66	46
Flime	0	30	85	52	29

group's ability to identify a stimulus-word properly was a function of its frequency of occurrence in English, while the younger group's identifications were influenced much less by familiarity with the stimulus-word. Table II lists the single words employed as stimuli in order of frequency of usage (based on the Thorndike and Lorge word count).[4] This table shows that all the older *Ss* identified the two most common stimulus-words properly at sometime during binaural or monaural stimulation, but that correct identification decreased with the less familiar words. Only 30% of the older group heard the nonsense-word (flime) correctly. Among the younger *Ss*, practically all managed to identify the real English words properly during the 3-min. tests. Even the nonsense-word was identified by 85% of the younger group.

Table II also shows the percentage of stimulation-time during which the word was heard correctly by those *Ss* reporting it. For each stimulus-word, the older people hearing it correctly kept with it a higher per-

[4] E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

centage of the time than did the younger *Ss*; there was a general tendency for the perception of less familiar words to be maintained for a smaller percentage of time by the older *Ss*. With the younger group, all English words were heard correctly about half the total time. Thus, the

TABLE III

REPORTS WHILE LISTENING TO SINGLE REPEATED WORDS

(The total number of forms reported is given for each stimulus-word followed by all forms reported by more than one *S*, the number reporting each form given in the preceding parentheses.)

TRESS

YOUNG: (72 forms). (19) tress; (11) press; (9) terez; (8) trez; (7) prez; (7) stress; (7) teress; (5) dress; (4) tressed; (3) tourist; (3) truss; (2) bread; (2) terass; (2) touress; (2) tred; (2) tresh; (2) trust.

OLD: (39 forms). (18) tress; (4) tressed; (3) dress; (2) rest-tress; (2) trash; (2) tread; (2) truss; (2) trust.

SEE

YOUNG: (73 forms). (20) see; (9) seeing; (8) pea; (6) thee [unvoiced 'th']; (5) bee; (5) tea; (4) fee; (3) dee; (2) bee-see; (2) being; (2) cease; (2) fee-see; (2) pink; (2) see-ink; (2) singing; (2) sink; (2) tea-see; (2) think.

OLD: (26 forms). (20) see; (3) three; (2) fee; (2) pea; (2) see-seed; (2) tea; (2) tea-see.

FLIME

YOUNG: (158 forms). (17) flime; (10) fly-'em; (8) fly; (7) slime; (6) climb; (6) fly-in; (6) flyed; (5) flying; (5) plime; (4) lime; (3) flibe; (3) flyer; (3) slide; (2) blime; (2) fline; (2) fly-in-flime; (2) flyumf; (2) fulime; (2) imply; (2) pline; (2) sly; (2) sulime; (2) wine.

OLD: (65 forms). (8) flyed; (8) slide; (7) fly; (6) flime; (6) flying; (6) slime; (4) climb; (3) flime-slime; (2) flying-flyed; (2) fried; (2) plied.

POLICE

YOUNG: (77 forms). (20) police; (15) please; (10) pleece; (10) peo-eece; (5) poe-eece; (4) police-please; (4) pulleys; (3) please-police; (2) oyce; (2) poyce; (2) boys; (2) pleece-police; (2) poe-iss; (2) poise; (2) pull-eece.

OLD: (27 forms). (20) police; (7) please; (4) please-police; (2) peace-police; (2) police-please; (2) priest.

TRICE

YOUNG: (101 forms). (20) trice; (9) tries; (7) price; (7) twice; (6) triced; (5) ter-ice; (5) try; (4) Christ; (4) right; (3) cry; (3) tries-trice; (3) try-ice; (2) krice; (2) estra; (2) her-eye; (2) prize; (2) ter-eyes; (2) thrice; (2) thrice-twice; (2) trice-tries; (2) try-ess; (2) try-est; (2) twiced; (2) twice-trice.

OLD: (44 forms). (15) trice; (6) tries-trice; (5) twice; (4) triced; (4) tries; (3) trice-tries; (3) tripe; (3) trite; (2) Christ; (2) Christ-twice; (2) trice-Christ; (2) trice-triced; (2) trice-twice; (2) tries-twice; (2) twice-trice.

most common word (*see*) and the least common English word (*trice*) were each heard 46% of the time. Only in the case of the nonsense-word *flime* does a clear difference appear (a drop to 29%).

Responses involving nonsense-words. The tendency of the older *Ss* to favor those words occurring frequently in English is indicated also by the low frequency of nonsense-responses for this group. Table III combines monaural and binaural trials and lists separtely all forms heard by two

or more *Ss* of the younger group and by two or more *Ss* of the older group. Conventional spelling is employed for the real English words, and the nonsense-responses are indicated by spelling chosen to afford a non-ambiguous version of the phonemes. It can be seen that many of the forms listed for the younger group involve nonsense-words. On the other hand, the only nonsense-forms listed for the aged group are those containing the stimulus-word *flime* itself, or the nonsense-word *flyed* (which suggests an incorrect past participle of the verb *to fly*).

While responses, especially for the younger group, were not restricted to real words, there is a rule followed by all responses for both age-groups. Only words containing phonemes in groupings permissible in English were reported. For example, a word such as *thlime* never was reported, the initial *thl* cluster not being found in English.

Discussion

Verbal transformations may be looked upon as involving two successive steps, organization and decay, one following the other again and again. First, there is organization of speech-sounds into a word or phrase. Next, there is a loss of organization with repetition. This decay of one perceptual form is accompanied by a return to the preceding process, *i.e.* organization of the speech-sounds into another word or phrase. The new form suffers the fate of its predecessor, and in its turn is replaced. These two processes have been investigated separately in the past.

Loss in meaning of words repeated by Ss. If, instead of listening passively to a repeated word, *S* is told to repeat a word himself, loss in meaning of the repeated word occurs.[5] Change in phonemes, such as observed for the verbal transformation-effect, is, however, restricted by the instructions to pronounce the same word again and again. Perceptual shifts to other forms do not occur with anything like the ease observed for listening to an externally produced repeating word.

Verbal organization of indistinct repeated speech. Skinner employed a recording which repeated a series of speech-sounds which were

---

[5] M. F. Bassett and C. J. Warne, On the lapse of verbal meaning with repetition, this JOURNAL, 30, 1919, 415-418; Molly Mason, Changes in the galvanic skin response accompanying reports of changes in meaning during oral repetition, *J. gen. Psychol.*, 25, 1941, 353-401; D.E.P. Smith and A. L. Raygor, Verbal satiation and personality, *J. abnorm. soc. Psychol.*, 52, 1956, 323-326; Michael Wertheimer and W. M. Gillis, Saturation and the rate of lapse of verbal meaning. *J. gen. Psychol.*, 59, 1958, 79-85; W. E. Lambert and L. A. Jakobovits, Verbal satiation and changes in the intensity of meaning, *J. exp. Psychol.*, 60, 1960, 376-383.

actually meaningless (a series of indistinct vowels without any consonants).[6] He found that, after listening to several repetitions of such a stimulus, *S*s reported hearing words and phrases. Skinner called his device the "verbal summator." After *S* responded, the recording was shut off. If stimulation had been continued, perceptual reorganizations probably would have been observed, as verbal transformations do occur with stimulation by a repeating word made indistinct by masking noise.[7]

Age-differences. It is known that the aged frequently have a decreased auditory acuity,[8] and the question naturally arises whether the age-differences found here may be explicable in terms of reduced loudness and clarity of the stimulus. Fortunately, data are available which enable us to answer this question.[9] It has been found that decreasing the intensity of a word repeated loudly and clearly to a faint but still intelligible level has no effect on the rate of illusory changes for young *S*s. It also has been found that if a repeating word is made quite unrecognizable by masking noise (the response bearing little resemblance to the stimulus), the rate of transitions drops, but still remains well above that obtained in the present experiment for the older group. It appears that resistance to illusory changes in the aged cannot be explained simply in terms of loss of auditory acuity with age. Central factors unrelated to intelligibility must be involved. The relative freedom of the aged from the perceptual instability of the young represents a case of more accurate auditory perception in the aged.

In the perceptual organization (and reorganization) of speech sounds, there was a surprising difference between the behavior of the two age-groups. The organizational units for the young group frequently appeared to be the English phoneme clusters rather than English words. Young *S*s invariably grouped speech sounds into those phoneme-groupings allowed in English, but organization frequently was not in the form of a real word, even when the actual stimulus was a common English word. By contrast, the aged tended to have real words as their organizational units, and usually reported common English words, even when the actual stimulus was a nonsense-word.

Visual analogues. The verbal transformation-effect was found through reasoning that listening to certain words repeated over and over might involve fluctuations analogous to those of the visual ambiguous figures

[6] B. F. Skinner, The verbal summator and a method for the study of latent speech, *J. Psychol.*, 2, 1936, 71-107.

[7] Warren, *op. cit.*, 249-258.

[8] A. D. Weiss, Sensory functions in J. E. Birren (ed.), *Handbook of Aging in the Individual*, 1959, 520-526.

[9] Warren, *op. cit.*, 249-258.

(*e.g.* the Necker cube).[10] It has been shown, however, that verbal transformations are not restricted to special configurations, and that they involve perceptual distortions without any parallel for visual ambiguous figures.[11] While verbal transformations do not seem related directly to visual ambiguous figures, they do appear to have other visual counterparts following equivalent perceptual rules.

The two processes which appear responsible for verbal transformations (dissolution of a particular perceptual organization with an unchanging pattern of stimulation and perceptual integration of unorganized sensory input) appear to exist not only for hearing, but for vision as well. The tendency for visual images to lose their organization and meaning with continued visual inspection has been described. As far back as 1907, Severance and Washburn reported that visual fixation of a printed word for a minute or two resulted in loss of meaning and a queer, unfamiliar appearance of the letters.[12] The tendency for non-representational visual forms to assume familiar meaningful shapes is illustrated by the Rorschach Test. A more striking example of this tendency may be observed with a completely unstructured homogeneous light-field (produced by placing half a celluloid table-tennis ball over the eye); *S*s viewing this homogeneous field may spontaneously report hallucinatory shapes and objects.[13]

Visual fluctuations involving perceptual distortions of non-ambiguous visual figures have been reported which seem analogous to the auditory verbal transformations. Marks instructed *S*s to observe a circle drawn on a card while fixating a small point.[14] Within a minute, most of them reported changes, some of which involved an elliptical or polyangular appearance of the circle, or even an incomplete figure with 'blotting out' of arcs. Recently, Pritchard, Heron, and Hebb have studied perception of a variety of visual figures viewed as stabilized retinal images.[15] Figures seen under these conditions were quite labile, and seemed to undergo a variety of changes suggestive of verbal transformations.

When the gaze is not fixed, but allowed to travel over a figure, the

[10] Warren and Gregory, *op. cit.*, 612-613.

[11] Warren, *op. cit.*, 249-258.

[12] Elizabeth Severance and M. F. Washburn, The loss of associative power in words after long fixation, this JOURNAL, 18, 1907, 182-186. See also V. J. Don and H. P. Weld, Lapse of meaning with visual fixation, this JOURNAL, 35, 1924, 446-450.

[13] J. E. Hochberg, William Triebel, and Gideon Seaman, Color adaptation under conditions of homogeneous visual stimulation (*Ganzfeld*), *J. exp. Psychol.*, 41, 1951, 153-159.

[14] M. R. Marks, Some phenomena attendant on long fixation, this JOURNAL, 62, 1949, 392-398.

[15] R. M. Pritchard, Woodburn Heron, and D. O. Hebb, Visual perception approached by the method of stabilized images, *Canadian J. Psychol.*, 14, 1960, 67-77.

continual change in neurological input permits the maintenance of a veridical perceptual organization. It may appear paradoxical, but nevertheless it seems to be that a stable perceptual organization cannot be maintained with an unchanging pattern of stimulation. Only by continually varying the sensory input—changing to neurologically different but perceptually equivalent patterns of stimulation—is it possible to maintain an accurate perceptual organization.

Almost all the work reported on fluctuation in visual perception has involved 'ambiguous' figures. With these figures (*e.g.* the Necker cube), changes occur during unrestricted eye-movements, and the transitions experienced represent plausible alternative interpretations of the stimulus with a minimal distortion of its actual configuration. The evidence for age-differences in number and in variety of changes while viewing ambiguous patterns is conflicting.[16] In any event, it seems clear that there are no age-differences of the magnitude observed with the verbal transformation-effect. It would be of interest to determine whether age-differences do exist for visual fluctuations more nearly analogous to verbal transformations (that is, for fluctuations while gazing with restricted eye-movements upon a non-ambiguous figure).

Summary

Young adults listening to a word or a short sentence repeated over and over on a loop of recorded tape experience sudden illusory changes in what they are hearing. These changes frequently involve considerable phonetic distortion of the clearly pronounced stimulus.

Marked age-differences in susceptibility to this 'verbal transformation effect' were found. *Ss* over 60 yr. of age exhibited a very much lower rate of verbal transformation and in many cases did not show this illusion at all. Perceptual organization of speech in the aged appeared to be influenced to a much greater degree by the frequency of occurrence of words in normal usage.

Verbal transformations for both age-groups were considered in terms of a mechanism involving two steps: (1) loss of perceptual organization with an unchanging pattern of stimulation; (2) perceptual (re)organization. Some relations between auditory and visual perception were discussed.

---

[16] For a summary of this evidence see A. T. Welford, *Ageing and Human Skill*, 1958, 169-172.

THE STIMULUS IN SERIAL VERBAL LEARNING

By Robert K. Young, University of Texas

The functional stimulus in serial verbal learning has not been identified specifically. Thus, when a series of items, A-B-C-D-E, etc., is learned as a serial list, it is difficult to determine whether the functional stimulus giving rise to Item D is C, all preceding items, the serial position of the item, all of these factors, or some other aspect of the situation.

In a previous study, a list of paired associates (*PA*-list) was presented *S* following learning of a serial (*Sr*) list.[1] Relatively little positive transfer occurred in learning the *PA*-list although it was constructed from the items of the *Sr*-list, in a manner which can be symbolized as A-B, B-C, C-D, D-E, etc. Thus, just as in the *Sr*-list, each item except the first and the last served as both a stimulus and as a response, and the pairings were consistent with the ordering of the items in the *Sr*-list in which A is considered a stimulus for B, B for C, C for D, etc. (Of course, during learning by paired associates, the pairs were never presented in a consistent order but always in different random orders.)

It would appear that if A is the specific and exclusive stimulus for B in the *Sr*-list, and if B is the stimulus for C, and so on, high positive transfer should obtain in the *PA*-learning since the A-B, B-C, C-D, etc., associations would have been learned in the *Sr*-list. As noted above, this is not the case. Furthermore, it is known that if this procedure is so reversed that *S* first learns a *PA*-list of the nature A-B, B-C, C-D, etc. (again randomly presented), and then learns an *Sr*-list in which the items are ordered A-B-C-D-E, etc., high positive transfer will occur.[2] In learning by *PA* it would appear that the particular verbal unit used as the nominal stimulus must inevitably serve also as the functional stimulus; there seems to be no other cue which *S* can use.

Since positive transfer should occur from *Sr*- to *PA*-learning as readily (and with the same magnitude) as positive transfer from *PA*- to *Sr*-learn-

---

* Received for publication September 10, 1960. This research was supported by Grant G-4850 from the National Science Foundation.

[1] R. K. Young, A comparison of two methods of learning serial associations, this Journal, 72, 1959, 554-559.

[2] Ernest Primoff, Backward and forward associations as an organized act in serial and in paired associate learning, *J. Psychol.*, 5, 1938, 375-395; Young, *op. cit.*, 554-559.

ing if the nominal stimulus is equally effective in both forms of learning, one possible conclusion from the facts as stated is that *S* is *not* learning associations between the specific verbal stimulus and the next item in the list in *Sr*-learning. Instead, he may be employing other features of the *Sr*-learning as the functional stimulus, *e.g.* serial position.

The above conclusion is exactly opposite to what may be called a specificity-hypothesis, according to which a functional stimulus for an item in a serial list *is* the item just preceding it. The present study tests certain implications of this hypothesis. If the implications are not borne out, the data may be used as evidence againt this hypothesis and in favor of some other hypothesis of serial learning.

If it is assumed that in learning a serial list A-B-C-D-E-, etc., A is the functional stimulus for B, B for C, C for D, etc., as per the specificity-hypothesis then when a *PA*-list is constructed (A-B, B-C, C-D, etc.) positive transfer to it should increase as degree of serial learning increases. Degree of serial learning is, therefore, varied in the present study to test this implication.

Assume that *S* has learned the *Sr*-list A-B-C-D-E, etc., and is then transferred to a *PA*-list consisting of pairs such as B-F, C-A, G-B, etc. If the specificity-hypothesis is valid, negative transfer should be expected in learning the *PA*-list constructed in such a fashion, because the specific associations developed in the *Sr*-learning (A-B, B-C, C-D, etc.) would be inappropriate for the *PA*-list. A test of this expectation from the specificity-hypothesis is also included in the present study.

METHOD

General. The experiment may be thought of as a 2 $\times$ 3 factorial design in which the columns represent three degrees of *Sr*-learning (given before *S* was transferred to a *PA*-list), and the rows two extremes in transfer; namely, a positive relation and a negative relation between the *Sr*- and *PA*-lists. The *Ss* in Group *P* (positive transfer) learned an *Sr*-list relevant to one of the three criteria of serial learning and then transferred to a *PA*-list so constructed from the *Sr*-list that, based on the specificity-hypothesis, positive transfer would be expected. The *Ss* of Group *N* (negative transfer) also learned an *Sr*-list to one of the three criteria, and then transferred to a *PA*-list constructed from the *Sr*-list with the items so arranged that negative transfer would be expected by the hypothesis. If the *Sr*-list is symbolized as A-B-C-D-E-F, etc., Group *P* followed the serial learning with a *PA*-list that would be symbolized as A-B, B-C, C-D, etc., while Group N learned the *PA*-list symbolized as B-F, C-A, G-B, etc. Thus the *PA*-list for both groups had the same stimuli and responses. The items for Group *P* were, however, paired in a manner consistent with the order of the *Sr*-list, while those for Group *N* were paired in a manner inconsistent with the order of the items in the *Sr*-list previously learned.

Specific procedures. Two basic groups of 24 *Ss* each were used, one group

serving under the *P*-conditions, the other under the *N*-conditions. The members of each group served under three condiitons, following a practice session, which were distinguished by the degree of learning given on the *Sr*-list before transferring to the *PA*-list. Under one condition there was no learning on the *Sr*-list; under a second, the *Sr*-list of nine items (eight actually learned since the first item was used as the anticipatory cue for the second) was learned until five out of the eight possible responses were given correctly on a single trial; and under the third condition, the *Sr*-list was learned to one perfect trial.

The intent of these procedures is to measure transfer on the *PA*-list as a function of degree of learning on the prior *Sr*-list. Differences in transfer (if present) are to be interpreted as being due to specific associations formed during the serial learning. It can be seen, however, that the learning of the *PA*-list might be influenced by differences in such general factors as 'warmup' and 'learning-to-learn,' resulting from varying degrees of prior serial learning. It was necessary, therefore, to make these general factors equivalent for all three conditions. To do this an

TABLE I
LISTS OF STIMULUS-MATERIALS

Sr-list: (Group *N*)	*Sr*-list: (Group *P*)	*PA*-list: (Both Groups)
entire	alike	solvent-unwell
icy	icy	entire-rustic
unwell	joyous	icy-joyous
rustic	solvent	unwell-clever
joyous	unwell	alike-icy
clever	clever	clever-entire
taboo	entire	joyous-solvent
alike	rustic	rustic-taboo
solvent	taboo	

irrelevant *Sr*-list (*ISL*) was presented before the relevant *Sr*-list (*RSL*) a sufficient number of times that, considering *ISL* and *RSL* combined, the amount of practice and warmup should be the same for all conditions at the time *S* was transferred to the *PA*-list. When *S* had *no* trials on the *RSL*, the *ISL*—irrelevant in the sense that no items from this list appeared in the *PA*-list—was given until *S* learned to one perfect trial. This condition is designated 8-0, indicating 8 correct out of the 8 possible on a single trial on the first or *ISL*, and zero trials on the *RSL*. For the medium degree of *RSL* learning (5 out of 8 correct) the *ISL* was first presented until 5 out of 8 were correct. This condition is designated 5-5. For the third degree of learning in which *S* learned the *RSL* to one perfect trial, *no* trials were given on the *ISL* and this condition is designated as 0-8. Both the *P*- and *N*-groups had exactly the same treatment in learning the *Sr*-lists. Since within each group all *Ss* served under all three degree-of-learning conditions, complete counterbalancing of conditions was used. It was, of course, also necessary to provide three different sets of lists, one for each condition. These sets of lists were also counter balanced, but independently of conditions.

The lists were composed of two-syllable adjectives taken from Melton.[3] Each *Sr*-list had nine items, each *PA*-list, eight pairs (see Table I). Intralist similarity

[3] E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 517-567.

and similarity between sets of lists were as low as careful inspection procedures permit. The *PA*-lists learned by the *P*- and *N*-groups were identical. The differences in the positive and negative characteristics obtaining between the *Sr*- and *PA*-lists were produced by variation in the ordering of the items in the *RSL*. For one set of items, this may be exemplified by presenting the *RSL* learned and the common *PA*-list.

It will be noted that while both groups learn the same *PA*-list, the relationship between the pairs and the items in the *Sr*-list differs for the two groups. For Group *N*, no pair in the *PA*-list represents two items which were adjacent in the *Sr*-list. For Group *P*, every pair represents adjacent items in the *Sr*-list and the pairing is in the direction appropriate to the *Sr*-list. That is, *alike* is the stimulus for *icy* in the *PA*-list which is exactly the case in the *Sr*-list. If the specificity holds—if *alike* is the functional as well as the nominal stimulus for *icy* in learning the *Sr*-list—positive transfer may be expected in learning the *PA*-list. So too (by the specificity-hypothesis) for Group *N*, since *entire* and *icy* should be associated during serial learning, negative transfer should result when *S* must learn the pair *entire-rustic* in the *PA*-list.

The learning of the *PA*-list for both groups under all conditions was carried to one perfect trial. Four different orders of presentation were used to minimize serial learning and thus emphasize learning in pairs. These orders were random but were subject to the restriction that two pairs containing the same item (once as stimulus and once as response) were never adjacent.

The *S*s were selected from introductory psychology courses and assigned to one of the two groups on a random basis. The lists were presented by a film-strip projector at a 2-sec. rate for the *Sr*-lists and a 2:2-sec. rate for the *PA*-lists. The intertrial interval for both kinds of lists was 6 sec. and approximately 1 min. elapsed between the learning of the two *Sr*-lists or between the learning of an *Sr*- and a *PA*-list. No more than one condition was given per day.

RESULTS

Control measures. On the practice day all *S*s learned the same *Sr*- and *PA*-lists. This allows a comparison between the 24 *S*s assigned to Group *P* and the 24 assigned to Group *N*. Since the t between trials to learn the *Sr*-list was 0.69, and that between trials to learn the *PA*-list was 0.18, the two groups can be considered essentially equivalent on learning ability.

For each group the total number of trials required in serial learning was determined for each condition. For Condition 0-8 this was simply the number of trials to learn the *RSL* to one perfect recitation. For Condition 8-0 the value was the number of trials to learn the *ISL*, and for Condition 5-5 the value was the total trials in learning both lists to a criterion of five out of eight correct. The values were compared by a Friedman two-way analysis of variance.[4] The χ^2_r for Group *P* was 2.04 and for Group *N*, 5.25. Since the χ^2_r needed for significance at the 5% level with 2 *df*.

[4] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956, 166-172.

is 5.99, neither of the Chi-squares obtained is significant. If, therefore, the number of trials on the two serial lists combined is used as an index of equality of learning-to-learn and warmup, the *S*s in each group learned a *PA*-list for each condition at the same level of practice and warmup on the average. Differences which occur in learning the *PA*-lists as a function of degree of relevant serial learning, must, thereby, be attributed to specific associative factors.

Trials. The mean numbers of trials to learn the *PA*-list to one perfect recitation for the three conditions for each group are shown in Table II. In the analyses of trials to learn, not a single *F* was significant. It can be seen from Table II that as degree of prior relevant serial learning increased, trials to learn decreased in Group *P* and increased in Group *N*. The *F* for interaction (Group × Condition) was 2.72, with 3.10 needed

TABLE II
MEAN TRIALS TO LEARN THE *PA*-LISTS

	Condition					
	8–0		5–5		0–8	
Group	M	σ_m	M	σ_m	M	σ_m
P	28.00	2.93	27.75	2.88	24.67	2.71
N	24.50	2.43	25.20	1.88	28.80	2.78

for significance at the 5% level with 2 and 84 *df*. The specificity-hypothesis implies that as degree of prior serial learning increases, the changes in learning for Groups *P* and *N* should be that shown by the means in Table II. Since the differences for interaction and those for conditions within experiments do not reach significance, it must be concluded that in terms of trials to reach one perfect recitation, the hypothesis is given weak support at best.

Transfer. In a previous experiment it was observed that positive transfer occurred early in *PA*-learning under conditions much like those of Group *P* in the present experiment.[5] This positive transfer was apparent until about half the items were learned, after which it decreased (relative to a control condition) and had disappeared by the time a criterion of one perfect recitation was reached. The present results for Group *P* show somewhat the same effect. To make the analysis, the mean number of trials required to attain a criterion of four correct responses was used. The means for this measure are shown in Table III. Because the variances have not been corrected for heterogeniety, the Friedman and Kruskal-Wallis analy-

[5] Young, *op. cit.*, 554-559.

ses of variance have been employed to test the significant of differences.[6] The difference beween Groups *P* and *N* was tested and $H = 10.07$ (with 1 *df.*, $p < 0.01$) indicating that Group *P* (positive transfer) reached a criterion of 4/8 earlier than did Group *N* (negative transfer). In addition, conditions in Group *P* were tested and $\chi^2_r = 15.02$ (with 2 *df.*, $p < 0.001$). The χ^2_r in Group *N* was found to equal 0.19 which is not significant. These results indicate that degree of prior *RSL*-learning influenced the transfer early in learning the *PA*-list for Group *P* but not for Group *N*. Even for Group *P* there is no clear gradient in transfer as a function of degree of serial learning. For Group *N* it must be concluded that there is no evidence at all for the specificity-hypothesis; a control condition (8-0) gave no better performance than the condition with the

TABLE III
MEAN TRIALS TO REACH A 4/8 CRITERION ON THE *PA*-LISTS

	Condition					
	8–0		5–5		0–8	
Group	M	σ_m	M	σ_m	M	σ_m
P	5.46	.63	2.88	.34	2.90	.45
N	4.83	.46	4.87	.49	5.58	.75

highest degree of *RSL*-learning, a condition which might be expected to produce negative transfer by the hypothesis.

The evidence evaluated thus far would give little support to the specificity-hypothesis. Only the positive transfer early in learning for Group *P* would support the notion and the differences, while significant statistically, are not of the order of magnitude to be expected if *S* merely continues to use the same associations in *PA*-learning as he had in *Sr*-learning.

Subsidiary experiment. It is conceivable that a higher degree of *RSL*-learning might produce results more in line with the specificity-hypothesis. To test this, as well as to make a further test of the hypothesis with a new sample of *Ss*, a subsidiary experiment was performed. In this experiment a group of 24 *Ss* served under each of three conditions: a positive transfer (*P*), a negative transfer (*N*), and a control condition. As in the major experiment, the *P*-condition employed a *PA*-list which had the same items and arrangements as the *Sr*-list it followed; the *N*-condition had a *PA*-list with the same items but with random pairing of the serial items and the control condition employed a *PA*-list which had items as unrelated as possible to those used in the serial list. In this subsidiary experiment the *PA*-lists were learned in a constant order with the conditions of transfer

[6] Siegel, *op. cit.*, 166-193.

counterbalanced. In all other respects the specific procedures of the major experiment were replicated with the exception that the criterion employed for serial learning was two successive perfect recitations under each condition.

The results of this second experiment were substantially the same as the first. Number of prior serial learning trials did not differ between conditions ($\chi^2_r = 0.77$) and trials to reach a one perfect criterion in the *PA*-learning did not vary as a function of the condition of transfer ($F = 1.95$ with 2 and 44 *df.*). As in the major experiment trials to reach a criterion of four correct responses was a function of transfer condition ($\chi^2_r = 7.52$, with 2 *df.*, $p < 0.05$) but the negative condition did not differ from the control ($\chi^2_r = 0.17$).

Further tests. It does not appear that when *S* learns a serial list A-B-C-D-E, etc., A is the functional stimulus for B, B for C, etc. Since, however, this conclusion may run against the commonly held conceptions of serial learning, further tests of the specificity-hypothesis were sought in the data. Several tests were performed on the data of both Groups *P* and *N*. The tests dealing with Group *P* will be presented first.

One of these tests deals with the bowed curve of serial-position. It may first be noted that a rather typical bowed-curve was found for the serial lists in the present experiments. For Group *P*, Condition 0-8, the mean number of correct responses per trial at each successive serial position was as follows: 0.91, 0.75, 0.67, 0.55, 0.45, 0.59, 0.65, and 0.68. Items near the first and near the last of the list are given more times correctly than those near the middle. It would seem that if specific associations are being transferred from the serial to the *PA*-list, some evidence for the bowed curve should be found in the learning of the *PA*-list for the *P*-conditions. For example, in the serial list A-B-C-D-E-F, etc., the A-B association should be the strongest at the end of learning and might, therefore, be expected to be given more times correctly in learning the *PA*-list if indeed A is the specific stimulus for B in the serial list. Likewise, a pair of items in the middle of the serial list, say D-E, having been given but few times correctly in serial learning, would be expected to show less transfer than A-B in the *PA*-list.

Differences in learning among pairs in the *PA*-list may occur for reasons other than differences attributable to serial learning, *e.g.* there may be differences in intrinsic difficulty of the pairs. To evaluate a transfer of the bowed-curve effect to *PA*-learning, a control or comparison condition must be employed. Condition 8-0 in Group *P* is the appropriate control, since the *S*s in this condition learned an irrelevant serial list. Therefore, the mean number of correct responses per trial for each pair was determined for Condition 8-0 and the pairs were ordered along a baseline in a manner which would have been appropriate had an *RSL* been learned previous to

the *PA*-list. This was compared with the results for Condition 0-8 which did have an *RSL*. The two curves are shown in Fig. 1. With the exception of the first and last serial positions, all points on both curves appear to have approximately the same value. The difference between the first points on each curve was compared by a sign-test and was found to be significant ($p = 0.01$).[7] No other differences between the two curves approach significance. Thus, following *RSL*-learning, which by the specificity-hypothesis

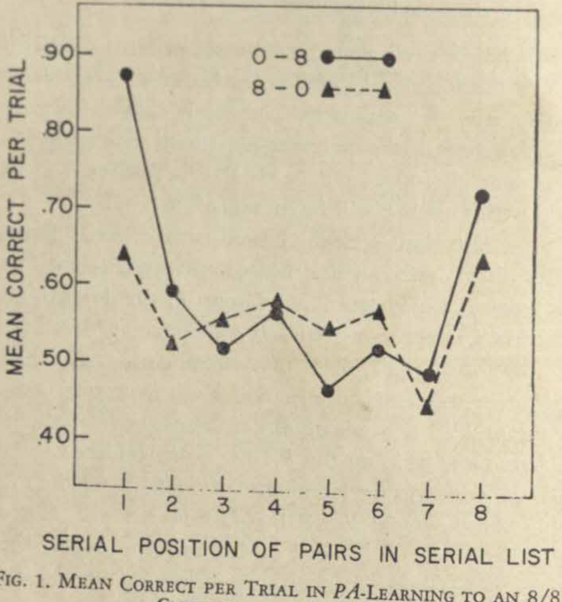


FIG. 1. MEAN CORRECT PER TRIAL IN *PA*-LEARNING TO AN 8/8 CRITERION FOR GROUP *P*

should produce high positive transfer, only the *PA*-pair which corresponded to the item in the first serial position was given correctly more often in the positive transfer condition (0-8) than in the control (8-0). Furthermore, a comparison of the shapes of the two curves gives little indication that transfer of the bowed-curve effect had occurred.

Similar analyses of the data of the subsidiary experiment were done and the *PA*-pair corresponding to the initial item in the *Sr*-list was correct more often in the *P*-condition than in the control ($p < 0.05$). No other evidence was found for transfer of the effect of serial position.

Since it has been observed that positive transfer does occur up to a

[7] Siegel, *op. cit.*, 68-75.

criterion of 4/8 in Condition 0-8, Group *P*, it might be assumed that the effect of serial position would transfer initially and then disappear when higher criteria of learning have been reached. The mean number of items correct per trial (up to a criterion of 4/8) in *PA*-learning is presented in Fig. 2 as a function of prior serial learning. As in the previous analysis, the pairs are ordered along a baseline which would have been appropriate had a prior *RSL* been learned. The *PA*-pairs from Conditions 8-0 and 0-8

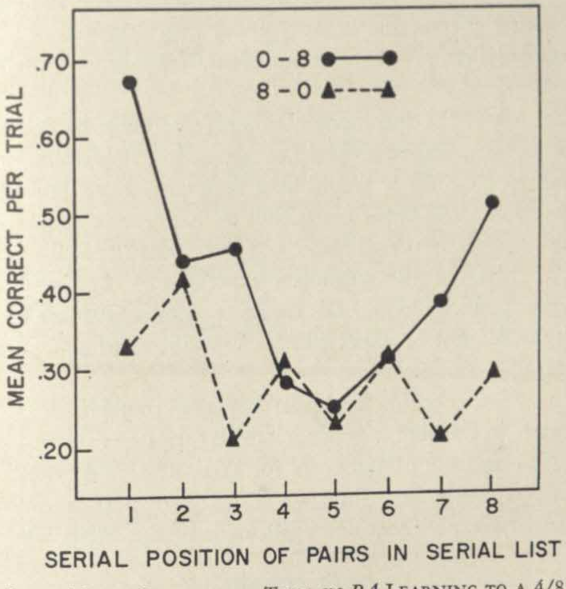


FIG. 2. MEAN CORRECT PER TRIAL IN *PA*-LEARNING TO A 4/8 CRITERION FOR GROUP *P*

in Group *P* were compared and again the first item was found to be given more often in Condition 0-8 than in Condition 8-0 ($p < 0.001$). No other pairs were found to differ from one another. There appears to be some evidence for transfer of serial-position effects in Fig. 2 with the *PA*-pairs corresponding to the first or last items in the serial list being given more often than the *PA*-pairs from the middle of the serial list. Similar results, however, were not found in the analysis of the trials to 4/8 criterion of the subsidiary experiment. None of the pairs in the *P*-condition was given significantly more often than the respective control pair and the plot of the serial position effects did not appear to be bowed.

The results of the tests of Group *P* give little indication that specific associations transfer from serial to *PA*-learning even though positive trans-

fer is observed early in the *PA*-learning of Group *P*. Further tests of the specificity-hypothesis may be made, however, from certain results of the negative-transfer conditions.

Following serial learning, the items were randomly rearranged in pairs in the *PA*-task to form a condition of negative transfer. Again, if the specificity-hypothesis is valid, the rearrangement of the serial items in the *PA*-task should result in negative transfer and while this did not occur in terms of trials to learn, the procedure employed may have resulted in negative transfer for specific items. Although no data will be presented, it can be reported that no evidence was found favoring a specificity-interpretation.

A further test of the specificity notion may be made by determining the number of incorrect responses given to stimuli in *PA*-learning which in the serial learning of Group *N* would have been correct. A sign-test indicated that *S*s gave more responses compatible with negative serial grouping in Condition 0-8 than in Condition 8-0 ($p < 0.05$). Similar results were found in the data of the subsidiary experiment. These data give some weight to the specificity-hypothesis since responses compatible with negative serial order would not be expected unless specific associations were formed in serial learning.

Since many experiments find that negative transfer effects disappear relatively early in learning, *e.g.* Underwood, *t*-tests were computed between Conditions 8-0 and 0-8 of Group *N* for Trials 1, 2, and 3.[8] None of these differences was significant. Similar analyses were performed on the data of the subsidiary experiment and again none of the *t*-tests was significant.

Discussion

The results of the present experiment give little support to the hypothesis that associations are formed to the nominal stimulus in *Sr*-learning. Such an hypothesis states that associations formed during *Sr*-learning would transfer to a *PA*-task composed of items learned during *Sr*-learning. The hypothesis would predict, further that positive transfer should obtain when the *PA*-list learned following *Sr*-learning, is composed of the same items and associations as were in the *Sr*-list while negative transfer should occur as a consequence of randomly re-pairing the serial items and constructing a *PA*-list from them. Furthermore, the specificity-hypothesis would state, in the case of the paradigm of positive transfer, that items which were often correct during *Sr*-learning would also be more often

[8] B. J. Underwood, Associative transfer in verbal learning as a function of response similarity and degree of first list learning, *J. exp. Psychol.*, 42, 1951, 44-53.

correct during *PA*-learning than would be the case with items which were less often correct in *Sr*-learning. On the other hand, the hypothesis would predict, in the case of the paradigm of negative transfer, that items most often correct during *Sr*-learning would show associative interference when transferred from *Sr*- to *PA*-learning. Finally, the predicted effects of positive and negative transfer should increase as degree of prior *Sr*-learning increases.

Little support for specificity has been found in the present study. In terms of trials to learn, transfer is not a function of degree of *RSL*-learning (except early in the *PA*-learning of Group *P*) and the analyses of transfer of specific associations gives little indication that the formation of associations during *PA*-learning is influenced by prior *Sr*-learning.

It may be argued that the present experiment is not a proper test of the specificity-hypothesis since the methods of presentation of *Sr*- and *PA*-tasks are different, *i.e.* since the two procedures are dissimilar—even though the mechanism of learning underlying both may be the same—no significant transfer-effects would be expected from one type of learning to the other. This objection appears questionable when viewed in the light of previous studies which found that high positive transfer obtained when the transfer from *PA*- to a *Sr*-task was measured.[9] It would appear that if the associations formed during *PA*-learning transfer readily to an *Sr*-task, then associations formed during *Sr*-learning should transfer as readily to a *PA*-task if the hypothesis that associations are formed to specific stimuli in *Sr*-learning is valid.

The results of the present experiments give little support to the specificity-hypothesis. While the data do not permit rejection of the hypothesis that associations are formed between the nominal stimulus and the response in serial learning, the data do indicate that such associations play a minor role at best. As noted above, many cues or combinations of cues may serve as the functional stimulus in serial learning and the results of the present experiments appear to cast doubt upon one possible explanation of the factors underlying serial learning—the association of a response to its nominal stimulus—but the data do not permit a choice among other alternate explanations.

SUMMARY

The present experiments test the hypothesis that associations are formed between the nominal stimulus and response in serial learning. This hypothesis, called the specificity-hypothesis, was tested by measuring transfer

[9] Primoff, *op. cit.*, 375-396; Young, *op. cit.*, 554-559.

from a serial to a paired associate (*PA*) list composed of the same items as were employed in the serial task. The *PA*-list employed in the present experiments differs from the usual *PA*-list in that each item served both as a stimulus, in one pair, and as a response, in another. Three conditions of transfer were employed. The first measured transfer from a serial to a *PA*-list when the items and associations were the same in both lists (A-B, A-B), the second when the items were the same in both lists but the associations were different (A-B, A-B jumbled) and the third was a control (A-B, C-D). The specificity-hypothesis would predict high positive transfer in the first paradigm, high negative in the second, and zero in the third.

The data were analyzed both in terms of trials to learn the *PA*-list and in terms of transfer of specific associations. The results indicate that positive transfer occurred only early in the *PA*-learning, and that negative transfer did not occur at any time during *PA*-learning. In addition the analyses of transfer of specific associations provided only weak support for the hypothesis. The results of the present experiments give little support to the hypothesis that associations are formed between the nominal stimulus and response in serial learning.

DISCRIMINATION IN AUDITORY AND VISUAL PATTERNS

By JO ANN S. KINNEY, U.S.N. Medical Research Laboratory

Within the realm of audition, there are relatively few data on the perception of complex sounds other than speech, and few laws on the perception of form and pattern with which to compare the laws of visual organization. Recent contributions to the field have been made by Miller and Heise in two experiments on auditory patterns.[1] In the first of these, they found that two alternating tones of different pitch will be heard as one pattern or trill if the separation in frequency between the tones is not too great. Beyond a certain size of separation, however, the trill breaks, and two unrelated tones are heard. Heise and Miller later experimented with other types of auditory pattern and found a similar phenomenon. A single variable tone, presented with others in a pattern, becomes isolated from the rest of the pattern when the frequency-separation is increased beyond a certain point. The "single-tone 'figure' is heard as an isolated 'pop' . . . with the onrushing steam of melodic pattern in the background."[2]

Temporal variation was not studied by Miller and Heise, but it is of obvious importance in auditory perception, since auditory stimuli are patterned in time. It does follow from their results, however, that if a tone is 'isolated' from the rest of an auditory pattern by a large frequency-separation, its temporal position with respect to the rest of the pattern should be more difficult to judge. Temporal discrimination therefore could be used to measure which tones are included in a perceptual group.

A further advantage resulting from the use of temporal discrimination as a measure of auditory patterning is that the results may be compared with data of visual patterning by means of the spectrographic transformation between vision and audition. Time is one of the major axes in this transformation, being comparable to the horizontal dimension of visual space, while pitch becomes the vertical dimension of space. There are a number of instances in the literature indicating that the spectrographic transformation preserves the patterning or organization of stimuli be-

* Received for publication June 15, 1960. The auditory portion of this paper is adapted from a Doctoral dissertation submitted to the University of Connecticut. The work was supported, in part, by the Haskins Laboratories in connection with Contract AF 19(604)-2285, which is monitored by Operational Applications Laboratory, AFCRC, Bedford, Mass. The author is indebted to Dr. Alvin M. Liberman who directed the research.

[1] G. A. Miller and G. A. Heise, Trill threshold, *J. acoust. Soc. Amer.*, 22, 1950, 637-638; An experimental study of auditory patterns, this JOURNAL, 64, 1951, 68-77.

[2] *Op. cit.*, this JOURNAL, 72.

tween the two senses. The first such indication came from the development of the Sound Spectrograph by Bell Telephone Laboratories and from the comparison of spectrographic and oscillographic displays of speech. Some individuals, after training, could read spectrograms of speech, a feat impossible with the oscillograph.[3]

The reverse of the Sound Spectrograph, the Pattern Playback developed by Haskins Laboratories, converts visual displays into sounds.[4] This instrument has been used successfully in determining for the first time the acoustic cues for speech, further attesting to the effectiveness of the transform in preserving the organization of the stimuli between the two senses.[5]

Studies of Licklider, Bindra, and Pollack on the intelligibility of speech,[6] and by Cooper, Liberman, and Borst on pattern perception,[7] also indicate the parallelism that exists between visual and auditory perception when the spectrographic transform is used. Heise and Miller conclude that their results with sound-patterning are in agreement with the laws of visual organization and that "the shape of the auditory pattern and the threshold for the integration of the variable tone into the pattern are approximately what one would expect from corresponding visual figures, if the frequency and time coordinates of the auditory figure are replaced by vertical and horizontal spatial coordinates, respectively."[8]

Other instances could be cited, the most obvious being that musical notation depends on this device and has been read successfully for centuries. It should be pointed out, however, that there is no indication that this transformation is the only one, or that it will work perfectly; it is simply, on the weight of the evidence, the most useful one we have at present.

This experiment then was designed, first, to provide quantitative data on the perception of temporal differences in auditory patterns and, secondly, to compare the auditory results with similar data from vision. Specifically, the effect of two major variables on auditory temporal discrimination was investigated. The frequency-separation of the elements or tones in the patterns was chosen as one variable, to see if the dependence of perceptual grouping upon pitch-relationships, found by Miller and Heise, could be repeated for temporal discrimination. The type of pat-

[3] R. K. Potter, Visible patterns of sound, *Science*, 102, 1945, 463-470.,

[4] F. S. Cooper, Spectrum analysis, *J. acoust. Soc. Amer.*, 22, 1950, 761-762; A. M. Liberman, Pierre Delattre, and F. S. Cooper, The role of selected stimulus-variables in the perception of the unvoiced stop consonants, this JOURNAL, 65, 1952, 497-516.

[5] F. S. Cooper, P. C. Delatttre, A. M. Liberman, J. M. Borst, and L. J. Gerstman, Some experiments on the perception of synthetic speech sounds, *J. acoust. Soc. Amer.* 24, 1952, 597-606; P. C. Delattre, F. S. Cooper, A. M. Liberman, and L. J. Gerstman, Speech synthesis as a research technique, *Proc. VIIth International Congress of Linguists* (1952), London, 1956, 545-561; A. M. Liberman, Some results of research on speech perception, *J. acoust. Soc. Amer.*, 29, 1957, 117-123.

[6] J. C. R. Licklider, D. Bindra, and I. Pollack, The intelligibility of rectangular speech-waves, this JOURNAL, 61, 1948, 1-20.

[7] Cooper, Liberman, and Borst, The interconversion of audible and visible patterns as a basis for research in the perception of speech, *Proc. nat. Acad. Sci.*, 37,1951, 318-325.

[8] Heise and Miller, *op. cit.*, this JOURNAL, 76.

tern, or the manner in which the tones are combined, was chosen as the second parameter, since it relates most obviously to form-perception in all modalities.

To obtain quantitative visual data for comparison, the spectrographic transform was used to convert the auditory patterns into visual patterns. Since there is as yet no evidence on the appropriate scale to use in such a conversion of axes, a linear transformation was arbitrarily chosen, with the result that the patterns were the same when the linear coördinates of frequency and time were replaced by linear dimensions of length. Thus, for visual experimentation, the elements of the patterns are rectangles, and the discriminative task is that of judging the spacing between

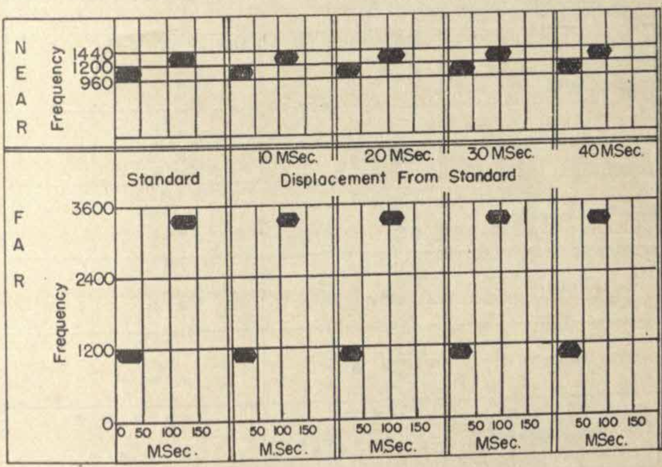


FIG. 1. SCALE-DRAWING OF THE TWO FREQUENCY-CONDITIONS OF THE BASIC ELEMENTS FOR THE STANDARD AND VARIABLES IN THE AUDITORY PATTERNS.

rectangles in the horizontal dimension. The two experimental variables are the amount of separation between elements in the vertical dimension and the type of pattern.

PART I: AUDITION

Apparatus. The stimuli were produced by the Pattern Playback, the instrument designed to convert spectrographic displays into sound. Spectrograms are painted in white on transparent material, placed on the moving apparatus, and passed through a modulated light beam which has a fundamental frequency of 120 ~ and 50 harmonics. The light reflected from the white paint is converted into the corresponding frequencies of sound by a phototube, amplifier, and speaker-system.

In this experiment, drawings were made of the desired patterns, sounds were generated from them by the Playback, and recordings of the sound were made on magnetic tape. The individual patterns were then placed in the desired orders for

presentation to the *Os* by marking and splicing the magnetic tape. The reels of magnetic tape were played for the *Os* on a Magnecorder and loud-speaker system in a room specially treated to absorb sound reflections.

Stimuli. The basic elements in each pattern of sound consisted of two tones of different frequency. The tones were each of 50 m.sec. duration and were separated in the standard condition by 50 m.sec. of silence. In the variable conditions, the silence was of 10, 20, 30, or 40 m.sec. duration.

Two degrees of frequency-separation were used; one small and the other large. The former will be referred to as the 'near' condition and the latter as the 'far.' In the near-condition, the two tones were of 1200 and 1440 ~; in the far-condition, they were 1200 and 3480 ~.

The individual tones were not pure but consisted of two adjacent harmonics of

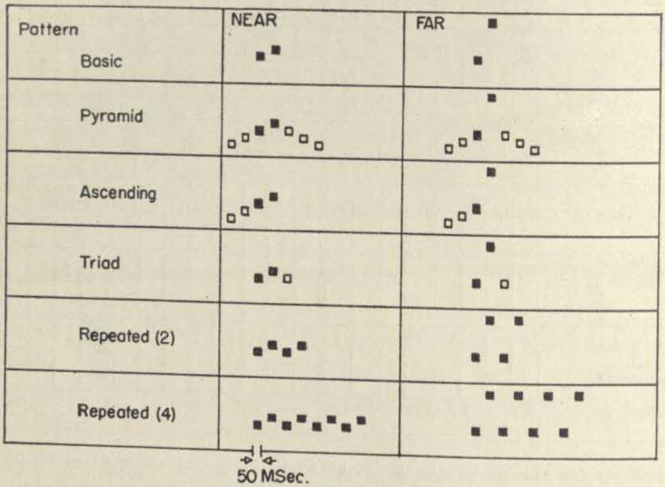


FIG. 2. SCHEMATIC DIAGRAM OF THE STIMULUS-PATTERNS

the 120 ~ fundamental; in specifying the frequency-level, the higher harmonic is cited. They started and ended gradually to avoid a burst of noise at onset and offset. Fig. 1 presents the two frequency-separations of the basic elements for the standard and variable conditions. This is a scale-drawing of the patterns actually used and painted for conversion into sounds. Two such drawings were made of each pattern. One of the patterns used in the experiment consisted of these basic elements presented alone. For each of the other patterns, exactly the same basic elements, drawn in the same way, either were imbedded in a series of other tones or were repeated a number of times by themselves.

Six different patterns were used. Fig. 2 shows a schematic diagram of these patterns, with the basic elements represented by filled squares and the other tones by open squares. Whenever tones other than the basic elements were added, the tones were of the same length and shape as the basic elements, and 50 m.sec. of silence was interposed between them.

This diagram presents only the standard condition of 50 m.sec. of silence between the basic elements. For each pattern, the variable conditions, with 10-40 m.sec. of silence between the basic elements, were exactly the same as those presented in Fig. 1. Thus, it was always the highest tone in each pattern that was displaced temporally away from the standard condition in which the spacings between tones always were equal; all other tones remained in the same position on the time-scale. This means that, for those patterns which had a tone following the basic elements, a given displacement from the standard condition narrowed the silence between basic elements and widened it following the variable tone. For example, at 40 m.sec. displacement, the silent-time was 10 m.sec. between the basic elements and 90 m.sec. between the displaced tone and the one following it.

Presentation of stimuli. The *ABX* method was used to measure the *O*'s ability to discriminate the temporal separation between the elements. This method presents three patterns in a series. The first two are different, one being the standard and the other, a variable; the third pattern in the series is exactly the same as either the first or the second. The *O*'s task, therefore, is to decide whether the third pattern, *X*, is like, the first *A*, or the second, *B*.

TABLE I

THE NUMBER OF *O*s TESTED ON EACH REEL

*O*s	No. sessions	Patterns presented: basic	pyramid	triad	ascending	repeated: two	four
Graduates	2	18	18	—	—	—	—
Undergrads	2	32	32	—	—	—	32
Undergrads	2	31	31	31	31	—	31
Undergrads	1	8	8	8	8	—	8
Undergrads	1	43	43	43	—	43	43
Columbia	1	6	6	6	6	—	6
Sailors	1	—	—	—	—	40	—
Total		138	138	88	45	83	120

The standard, with 50 m.sec. of silence between the basic elements, was compared with each of the variables that had 10-40 m.sec. of silence between elements. All possible orders were used for each comparison—*ABB*, *ABA*, *BAA*, and *BAB*. To control for the possibility that variations in the drawing or recording of the stimuli might serve as clues to the correct answer, the two identical temporal variations in each *ABX* series always were made from different drawings and recordings of the same pattern. A complete test for any one pattern consisted of all four orders for each of the four displacements, or 16 *ABX* series. These 16 series were placed on one reel of magnetic tape with the order of presentations randomized. Two frequency-conditions for each of the 6 patterns gave 12 reels.

The *O* was told that the difference between the patterns was of a temporal or rhythmical nature and to listen carefully for this aspect. He was further told that some of the differences might be difficult to detect and asked, where he was unsure, to guess. Each *O* was given data sheets on which to write down his answers.

These sheets were scored for the number of correct judgments at each of the temporal displacements.

The *Os* were, for the most part, undergraduates at the University of Connecticut, but they also included some graduate students from Columbia University, and enlisted men of the U. S. Navy. They were tested in small groups of less than 20 and heard as many reels as possible during the allotted testing time. All reels were not played to all *Os*, since an individual session, at most 45 min. long, was never long enough to encompass all of the variations; but the same *Os* always were tested with both the near and far frequency-conditions of a single pattern. The order of presentation of reels given within a single session was randomized, and each small

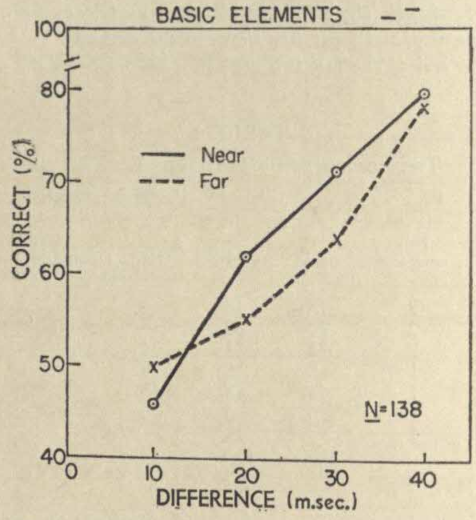


FIG. 3. RESULTS FOR THE TWO FREQUENCY-CONDITIONS OF THE BASIC ELEMENTS

group of *Os* was assigned a different order. The *Os* usually were asked to return for a second session a week later, and the same reels were repeated, in a different order, for those who did return. Table I gives all the information on the various patterns to which each group of *Os* was exposed and the total number of *Os* who heard each reel.

Results. The results obtained with the two frequency-separations of the basic elements are presented in Fig. 3, in terms of the average percentage correct for 138 *Os* at each temporal displacement. Normal psychophysical functions are found relating discrimination to the size of the difference to be discriminated. Both curves start with a chance-level of discrimination at 10 m.sec. and show increasingly larger percentages as the length of the temporal displacement is increased. The liminal values for both curves,

at the 75% point, are 30 -40 m.sec. The near-separation yields better discrimination throughout the curve except at the 10-m.sec. interval. This same increase in discrimination in the near-condition is found in the results for all of the other patterns. These data are given in Figs. 4 and 5;

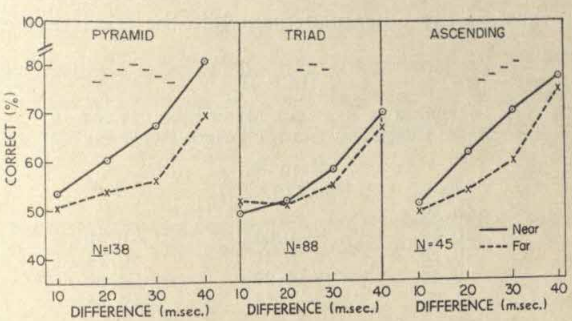


FIG. 4. RESULTS FOR THE TWO FREQUENCY-CONDITIONS OF THE PYRAMID, TRIAD, AND ASCENDING PATTERNS

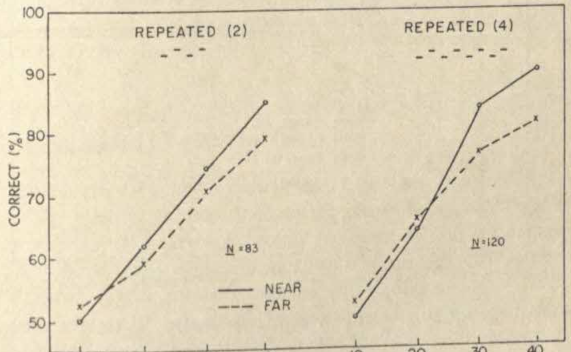


FIG. 5. RESULTS FOR THE TWO FREQUENCY-CONDITIONS OF THE REPEATED PATTERNS

wherever the results are above the chance-level, the near-condition yields better discrimination than the far.

It is obvious also from these data that discrimination varies with the type of pattern presented even where the actual temporal differences to be discriminated are identical. As a further check on this finding, comparisons were made only for *O*s who responded to all the reels in question; the purpose, of course, was to eliminate the effects of individual differences

in sensitivity. In the previous comparison between frequency-separations, each *O* always was given both frequencies of a single pattern, which meant that individual differences already were eliminated. The results of these comparisons between patterns revealed consistent differences for all groups. The repeated patterns always showed discriminative functions superior to those of the other patterns, while the triad gave the poorest

TABLE II

THE DIFFERENCES BETWEEN MEAN RESULTS FOR THE BASIC ELEMENTS AND THE OTHER PATTERNS

Pattern	Mean	Mean diff. (Basic—Other)	*t*	P
Basic	10.39			
Pyramid	10.43	−0.04	−0.16	—
Ascending	10.18	0.21	0.74	—
Triad	9.10	1.29	4.80	1%
Repeated	11.91	−1.52	−4.95	1%

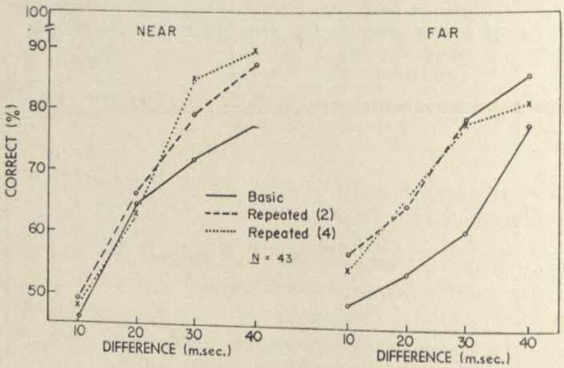


FIG. 6. EFFECT OF REPEATING THE BASIC ELEMENTS ON DISCRIMINATION

functions. The *t*-tests in Table II are given as an example of these differences; comparisons are made between the basic elements and four other patterns for a group of 31 *O*s, all of whom responded to all five patterns.

Data on the repeated patterns for a group of 43 *O*s, all of whom had the basic elements and two and four repetitions of the basic elements are plotted in Fig. 6. In both the near- and the far-conditions, there is better discrimination for the repeated patterns than for the basic elements alone. It should be noted also that the number of repetitions makes little difference in the discriminative functions, the curves for two repetitions and

for four repetitions intertwining at both near- and far-frequencies. Table III gives the results of *t*-tests performed on these data. No significant differences exist between two and four repetitions of the basic elements, while the differences between the basic elements and its repetitions all are significant.

PART II: VISION

The visual patterns used in this experiment were drawn in black India ink on 3 × 5-in. white cards. For presentation to the *O*s, they were placed in a holder against a large, white, Bainbridge-board screen. The screen

TABLE III

DIFFERENCES BETWEEN THE MEAN RESULTS FOR THE BASIC ELEMENTS, TWO, AND FOUR REPETITIONS

Frequency	Patterns	Mean diff.	t	P
Near	Basic *vs.* two	0.86	1.82	5%
	Basic *vs.* four	1.00	2.23	1%
	Two *vs.* four	−0.12	−0.34	—
Far	Basic *vs.* two	1.84	3.60	1%
	Basic *vs.* four	1.56	3.45	1%
	Two *vs.* four	0.28	0.92	—

* The value given is for a one-tailed test, on the assumption that the repetitions always should be better than the elements alone.

was placed at the end of a 30-ft. alley which was painted a neutral grey and evenly illuminated to a level of 14 ft.-c.

Stimuli. The basic elements in each visual pattern were two solid rectangles, ¼ × ⅛ in. The horizontal space between the two rectangles varied from ¼ — ⅛ in. in 1/32-in. steps. At the viewing distance of 26 ft. 10 in., these steps gave separations of 2⅔, 2⅓, 2, 1⅔, and 1⅓ min. of visual angle. This variation in the horizontal dimension was used to measure discrimination.

Two degrees of spatial separation in the vertical dimension were used. In one, the 'near-condition,' the top of the first rectangle was in line with the bottom of the second. In the other, the 'far-condition,' 1 1/16 in. or 11.3 min. separated the two rectangles. A scale-drawing of the variations in the spacing of the basic elements is given in Fig. 7.

Five patterns were used: the basic elements, two and four repetitions of the basic elements, the triad, and the pyramid. The schematic drawing of the auditory patterns in Fig. 2 is equally applicable to the visual patterns, except that the 50-m.sec. interval is ¼ in. in the visual patterns. When other rectangles were added to the basic elements, a constant, horizontal spacing of ¼ in. was used, with only the top rectangle changed in spacing in relation to the others. Thus, as the top rectangle was moved to the left, the distance between it and the rectangle on its right was increased proportionately. Each of the variations of the patterns was drawn on its own 3 × 5-in. card. Two drawings were made of each variation.

Presentation of stimuli. The *ABX* method was used as in the auditory part. *O* was presented with three cards in succession and asked whether the third pattern was like the first or the second. The three cards varied in the amount of horizontal displacement between the basic elements.

Preliminary observations showed that *O*'s discrimination of the difference in spacing was extremely acute. No standard, therefore, was used, as was the case in the auditory problem, but rather each variation in spacing was combined with every other variation to obtain as fine a measure as possible. Thus, the 8/32-in. spacing between elements was combined in *ABX* series with 7/32 in.; 7/32 in. was combined with the 6/32 in.; 6/32 in. with 5/32 in.; and so forth. All such combinations, in which the difference to be discriminated was 1/32 in. (1/3 min.), are called the one-step differences. All combinations of 2/32-in. (2/3 min.), 3/32-in. (1 min.), and 4/32-in.

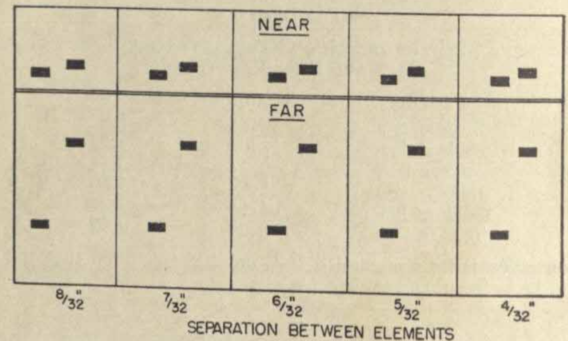


FIG. 7. SCALE-DRAWING OF THE BASIC ELEMENTS USED IN THE VISUAL PATTERNS

1⅓ min.) differences also were used. Sixteen presentations of each step-difference of a given pattern were made in random order in one session. Two sessions were given on each pattern, yielding 32 judgments by each *O* for each of the one-, two-, three-, and four-step differences.

At the beginning of each session, the series of five variations of the pattern were shown simultaneously to the *O* that he might see the differences to be discriminated. The cards then were removed and thereafter presented singly. Each card was exposed for 5 sec. with about 0.5 sec. between exposures in the same *ABX* series. The *O* was instructed to guess if he could not tell the difference between the patterns.

Four *O*s were used, all of whom had normal visual acuity of about 20/15 as measured by the Snellen Letter Chart. The order in which the patterns were presented was randomized except that each *O* completed all the patterns before repeating them the second time.

Results. The average results of the four *S*s are presented in Fig. 8 with the percentage of correct responses plotted as a function of the visual angle to be discriminated. The effect of the vertical spatial separation on the discriminative functions is shown separately for each pattern, the near-

separation always resulting in better discrimination than the far. The same superiority of near over far was found in the curves of each individual, although there were individual differences in the over-all level of discrimination.

The curves for the basic elements presented alone show that a difference of ⅓ min. of visual angle could not be discriminated as different by the *O*s. With larger visual angles, performance improves, the 75%

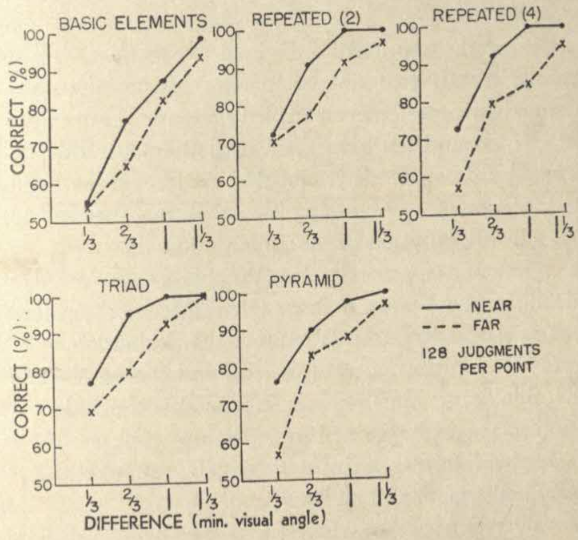


FIG. 8. RESULTS FOR THE TWO DEGREES OF VERTICAL SEPARATION IN EACH OF THE PATTERNS

point falling at about ⅔ min. for the near-condition and at about ⅚ min. for the far-condition.

Further inspection of Fig. 8 shows that, both in the near- and far-condition, repetition improves discrimination, although four repetitions are no more helpful than two. The triad and the pyramid yield better discrimination than the basic elements, although the pyramid is no better than the triad.

DISCUSSION

The effect of the frequency-separation of tones upon perception of sound-patterns, found by Miller and Heise, has been reproduced here under quite different conditions. It will be remembered that Miller and

Heise found a large frequency-separation between one of the tones in a pattern caused it to 'pop' out of the pattern.[9] It was pointed out previously that this effect should manifest itself in poorer temporal discrimination because the time-relations between an 'isolated' tone and the rest of the pattern should be more difficult to judge than the relations between tones all of which are in a single pattern. This expectation has been confirmed: the results show that temporal discrimination is poorer for the tone separated from the rest of the pattern by a greater difference in frequency.

A comparison of the results of the visual experiment with those from the auditory one reveals that one of the two major effects studied, that of the separation between elements, yielded the same result in both modalities. For both vision and audition, increasing the separation between the elements resulted in poorer discrimination under every condition. This then is direct support for the supposition that the spectrographic transform preserves the organization of the stimuli.

When the effect of the second major variable is considered, the results are not so definitive. One point is in complete agreement; patterns formed by repetitions of the basic elements improve discrimination in both senses, the differences between the repeated patterns and the basic elements alone being highly significant. Furthermore, it may be concluded that, while there is an advantage to be gained by repeating the basic elements once, there is no further advantage in increasing the number of repetitions. This is true for both vision and audition.

Further consideration of the effects of pattern-types, however, reveals an interesting discrepancy. The auditory triad yielded the poorest discriminative functions, while the visual triad was superior to the other patterns. Again, the basic elements ranked second to the repeated patterns in audition, while in vision this pattern was consistently the poorest.

This result indicates that the importance of some factor has been underestimated in the visual system. A possible consideration is that symmetry of pattern plays an important part in visual discrimination but not in auditory. Evidence for this supposition is provided by the *O*s who participated in the visual work. Each of them reported individually and spontaneously that judgments of the repeated patterns, under the far-condition, were impossible by trying to look at all of the elements. They used instead only the first three elements and made their judgments on this basis, the problem now being one of symmetry, and identical to the

---

[9] Heise and Miller, *op. cit.*, this JOURNAL, 72.

far triad, except that they had to ignore the unwanted elements. Further study is required to determine whether a comparable factor can be found in audition by varying the conversion-factors on the axes of the spectrographic transform, or whether this is a factor specific to vision for which there is no transformation comparable in audition.

SUMMARY

The effect of two variables on temporal discrimination in auditory patterns and on spatial discrimination in visual patterns was measured. The two variables were the type of pattern, or the manner in which the elements were combined, and the degree of separation between elements. The patterns were transformed between the two senses by making frequency in audition comparable to the vertical dimension of visual space, and time in audition comparable to the horizontal dimension of visual space.

Temporal discrimination in auditory patterns was better when the elements in the pattern were of relatively near frequencies than when they were widely separated. This outcome was in accord with the visual result of better spatial discrimination when the elements were close together in the vertical dimension than when they were far apart.

Temporal discrimination also varied with the type of pattern in which the discrimination was imbedded. One repetition of the basic elements improved discrimination, but there was no further increase with further repetitions. The same result was found for visual (spatial) discrimination. The additional elements provided in the non-repeated patterns, as the triad and pyramid, did not aid discrimination in audition, but did improve it in vision.

PROBABILITY-MATCHING IN THE FISH

By Erika R. Behrend and M. E. Bitterman, Bryn Mawr College

The only animal other than man which has yielded unequivocal evidence of probability-matching is the African mouthbreeder, *Tilapia macrocephala*.[1] In the first of two exploratory investigations reported several years ago, a small number of these fish were trained on a simultaneous horizontal-vertical discrimination with response to one of the stimuli reinforced on a random 70% of trials and response to the other reinforced on the remaining 30% of trials. This distribution of reinforcements was ensured by the use of a 'guidance' procedure: if on any trial the unreinforced stimulus was chosen initially, it was removed, and *S* was permitted to earn a reinforcement for response to the other. Under these conditions, each animal in the group developed a stable tendency to choose the more frequently reinforced stimulus approximately 70% of the time. In the second experiment, the same animals were trained with two gray stimuli on a 70:30 spatial problem. Again, matching developed, although the preference for the more frequently reinforced position shifted rapidly from 70% to 100% when guidance was abandoned in favor of a pure noncorrectional method. Analogous experiments with rats and with monkeys have yielded no indication of matching. These animals maximize; that is, they tend to choose the more frequently reinforced stimulus almost 100% of the time.[2] The experiments reported in the present paper were designed to provide further information on the course of probability-learning in the mouthbreeder and the conditions under which matching occurs.

A secondary concern of the present work was with the course of habit-reversal after inconsistent reinforcement. The results for mammals are

* Received for publication July 15, 1960. This work was supported by Grant M-2857 from the National Institute of Mental Health.

[1] M. E. Bitterman, Jerome Wodinsky, and D. K. Candland, Some comparative psychology, this Journal, 71, 1958, 94-110.

[2] Bitterman, Wodinsky, and Candland, *op. cit.*, 103-108; Allen Parducci and James Polt, Correction *vs.* noncorrection with changing reinforcement schedules, *J. comp. physiol. Psychol.*, 51, 1958, 492-495; W. A. Wilson, Jr., and A. R. Rollin, Two-choice behavior of rhesus monkeys in a noncontingent situation, *J. exp. Psychol.*, 58, 1959, 174-180; Wilson, Two-choice behavior of monkeys, *ibid.*, 59, 1960, 207-208; D. R. Meyer, The effects of differential probabilities of reinforcement on discrimination learning by monkeys, *J. comp. physiol. Psychol.*, 53, 1960, 173-175.

suggestive of the paradoxical effect of partial reinforcement on resistance to extinction; for example, rats reverse less rapidly after 67:0 than after 100:0 training,[3] and monkeys reverse less rapidly after 60:40 than after 60:0 training.[4] Since mouthbreeders do not show the paradoxical effect in simple instrumental training—their initial resistance to extinction is *reduced* by partial reinforcement[5]—there was reason to be curious about the readiness with which they would reverse a preference established with inconsistent reinforcement.

Method

Subjects. The 16 animals used were not experimentally naïve. All had had rather extensive training in a simple instrumental (single-target) situation and some limited extinction-experience.

Apparatus. The technique employed was an extension of that described by Longo and Bitterman.[6] Two interchangeable targets of light metal—in this case, one black and one white—were introduced into *S*'s individual 2-gal. living tank, as shown in Fig. 1. The targets were mounted on light rods inserted into the needle-holders of crystal phonograph-cartridges. The outputs of these cartridges were amplified and used to operate a set of relays. (Such a system is extremely sensitive to contact of fish and target, which, of course, is precisely what it was designed to detect.) The reinforcement was a pellet of food (Aronson's mixture) dropped into the water by an automatic feeder. Either of the two targets, or both of them together, could be introduced into the water, or withdrawn from it, by *E*, as indicated in Fig. 1. The fish was brought to the experimental situation in its living tank, the long sides and the back of which were painted in such a way that they admitted only diffuse light. The front of the tank, which was of clear glass, was set before a gray-painted background against which the targets were seen.

Procedure. After several days of pretraining in the new situation, during which the animals learned to strike at both targets readily, there were two days of 50:50 training with guidance, during which the preferences of the animals were established. Then the animals were divided into two groups of 8 *Ss* each, matched for the direction and strength of preference, and experimental training was begun. The problem was a confounded brightness-position discrimination (the black target always in one position and the white target always in the other). With minor exceptions (to be noted later) there were 20 massed trials per day. In some stages

[3] J. H. Grosslight, J. F. Hall, and Winfield, Reinforcement schedules in habit reversal—a confirmation, *J. exp. Psychol.*, 48, 1954, 173-174.

[4] C. B. Elam and D. W. Tyler, Reversal-learning following partial reinforcement, this JOURNAL, 71, 1958, 583-586.

[5] Jerome Wodinsky and M. E. Bitterman, Partial reinforcement in the fish, this JOURNAL, 72, 1959, 184-199; Resistance to extinction in the fish after extensive training with partial reinforcement, this JOURNAL, 73, 1960, 429-434. Nicholas Longo and M. E. Bitterman, The effect of partial reinforcement with spaced practice on resistance to extinction in the fish, *J. comp. physiol. Psychol.*, 53, 1960, 169-172.

[6] Longo and Bitterman, Improved apparatus for the study of learning in fish, this JOURNAL, 72, 1959, 616-620.

of training, guidance was used and in other stages it was not. Guidance involved the withdrawal of both targets after an unreinforced response and the reintroduction only of the target which was positive on that trial. In the final stage of training, guided choices of this kind sometimes were scheduled quite independently of erroneous choices.

RESULTS

Experiment I. In the first experiment, both groups were trained, with guidance, against the preferences demonstrated in pretraining.

One group was put on the 100:0 problem, and the second group was put on the 70:30 problem. For the 70:30 group, reinforcements of the minority stimulus were

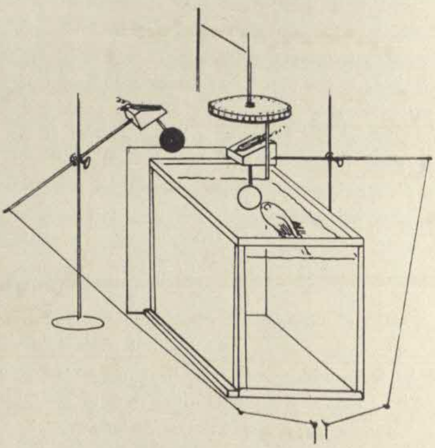


FIG. 1. DIAGRAM OF THE EXPERIMENTAL SITUATION

randomized over blocks of 10 trials, with the restrictions (1) that there could be no more than two such reinforcements in succession, and (2) that there should be at least one in each block of five trials. On Day 25, both groups were shifted to the 0:100 problem, the former minority stimulus now being consistently reinforced. Training on this problem continued through Day 38.

Plotted in Fig. 2 are the daily percentages of preference for the stimulus more frequently reinforced at the outset of training. The curves begin at a low level, because the animals were trained against their original preferences, and then rise rapidly, the 100:0 curve to a very high level, and the 70:30 curve to the 70% level. The matching shown by the 70:30 curve is not merely a group effect, but an individual phenomenon; the preferences of individual *S*s averaged over Days 4-24 range from 59-76% with a mean of 68%. The individual preferences for the 100:0

group averaged over Days 7-24 range from 93-99% with a mean of 96%.

Shifted to the 0:100 problem, both groups reversed rapidly. After the first day or two on the new problem, their performance was almost identical. Since the curve for the 100:0 group begins at a higher level, its initial rate of change is greater, but there is no hint of cross-over. Unfortunately, there is no directly analogous mammalian experiment which may be compared with this one, but extrapolation from available mammalian data suggests that the results of a directly analogous experiment would be rather different than those pictured here.

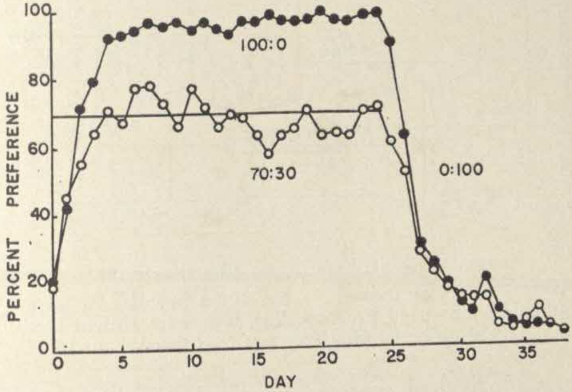


FIG. 2. PERFORMANCE IN EXPERIMENT I

(One group was trained on the 100:0 problem, a second group on the 70:30 problem, then both were shifted to the 0:100 problem. Guidance was used throughout.)

Experiment II. Earlier findings for the 70:30 problem having been confirmed in the first experiment, the next step was to extend the search for matching to other probability-ratios.

In this experiment, the original groups of fish were split to form two new groups, equated for original group-membership and for performance in the first stage of the experiment. One of the new groups was put on the 20:80 problem and the second on the 40:60 problem for 22 days, after which both groups were shifted to the 50:50 problem for 17 days. Guidance was used throughout. In the 20:80 training, one reinforcement of the minority stimulus was scheduled at random for each block of five trials. In the 40:60 training, there were two such reinforcements in each block of five trials and no more than two in succession. In the 50:50 training, the positive stimulus was designated according to selected Gellermann-orders.

The performance of the animals in the second experiment is plotted

in Fig. 3.[7] The points for Day 38 show the performance of the two groups on the last day of 0:100 training. Thereafter, the two curves separate, approaching different asymptotes in negatively accelerated fashion. The 20:80 group tends at asymptote to choose the minority stimulus approximately 20% of the time. Over Days 48-60, the individual preferences range from 11-23% with a mean of 19%. The curve for the 40:60 group tends somewhat to undershoot the 40% level, the mean for Days 48-60 being 34% and individual scores ranging from 30-48%. There is no overlapping whatsoever between the two groups, either in daily means

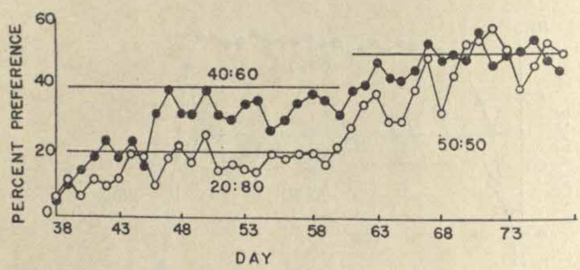


FIG. 3. PERFORMANCE IN EXPERIMENT II
(One group was trained on the 20:80 problem, a second group on the 40:60 problem, then both were shifted to the 50:50 problem. Guidance was used throughout.)

or in terms of individual preference-levels. It should be noted that an increase in preference for the minority stimulus as animals are shifted from 0:100 to 20:80 or 40:60 means a decrease in the number of initial reinforcements, although (with guidance) the total number of reinforcements remains the same.

During the period of 50:50 training, the mean preferences of both groups shift once more, this time to the 50% level. The change in ratio of reinforcement is reflected in the performance of each animal in each of the two groups. Individual preferences averaged over Days 67-77 range from 31-69%. Confronted with insoluble problems of this sort, mammals develop strong preferences for one or the other of the stimuli. Group curves may run along at the 50% level, but only because the two stimuli are fixated by equal numbers of *Ss*.[8]

In Fig. 4, the average deviation of individual asymptotic preferences

---

[7] The curves are based on the data for 7 *Ss* in each group, 2 *Ss* having been lost in the course of the experiment.

[8] Meyer, *op. cit.*, 174.

about asymptotic preferences for the group are plotted for the probability-ratios studied in Experiments I and II. As the curve shows, variability decreases monotonically as the ratio deviates from 50:50.

Experiment III. For the third experiment, two new groups of animals were constituted, matched for performance in the second phase of Experiment II, and both groups now were trained, *without guidance,* against

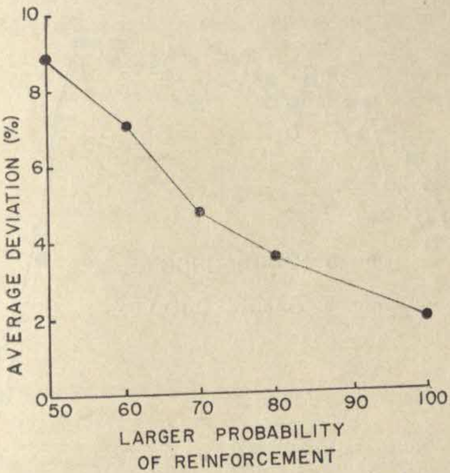


FIG. 4. PRECISION OF MATCHING AS A FUNCTION OF PROBABILITY-RATIO

(The measure of precision is the average deviation of the preferences of individual *Ss* from the mean preference of the group. The curve is based on the data of the first two experiments.)

whatever preference was manifested in that performance, one on the 100:0 problem and one on the 70:30 problem.

The purpose of the work was to check on the earlier conclusion of Bitterman, Wodinsky, and Candland that guidance is essential for matching.[9] After 21 days, both groups were shifted to the 0:100 problem. With the use of a noncorrection method, the amount of food earned by *S* depended on its choices; total daily intake was equated in supplementary feedings, the number of pellets given each *S* being equal to the number of errors plus two.

The performance of the two groups is plotted in Fig. 5. The first point is for the last day of the previous experiment, and it is below the chance-

[9] Bitterman, Wodinsky, and Candland, *op. cit.*, 106.

level because the initially preferred stimulus was the minority stimulus. Thereafter, both curves rise in negatively accelerated fashion to a high level. That is, the 70:30 group maximizes rather than matches, a result which confirms the earlier conclusion. Before the point of shift, the performances of the two groups are, in fact, statistically indistinguishable. After the shift, too, the two groups perform in comparable fashion; that is, inconsistent reinforcement neither increases nor decreases the difficulty

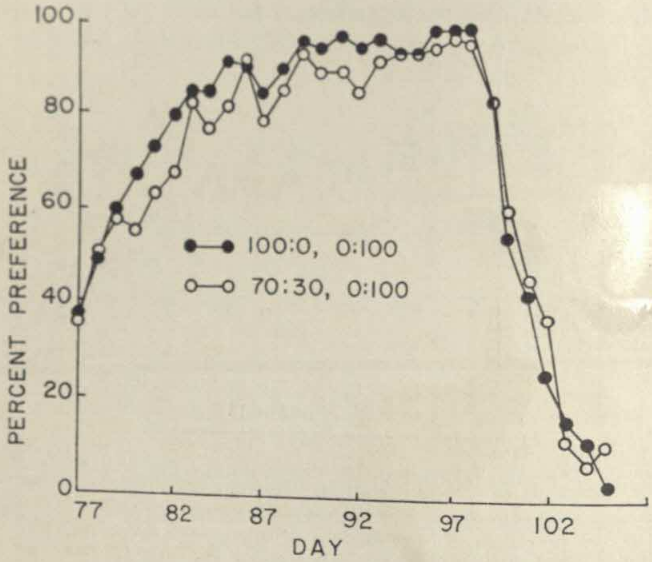


FIG. 5. PERFORMANCE IN EXPERIMENT III
(One group was trained on the 100:0 problem, a second group on the 70:30 problem, without guidance after error in either case. Then both were shifted to the 0:100 problem.)

of habit-reversal. Again, a directly analogous mammalian experiment is not available, but existing data suggests that such an experiment would yield rather different results.

Experiment IV. Matching appears when guidance is used but not in simple, noncorrectional training. Will any procedure which ensures a certain distribution of reinforcements between two stimuli produce a corresponding distribution of choices? The fourth experiment was designed to answer this question.

In the first stage of the experiment, a procedure used by Ramond in work with the rat was adopted.[10] Each animal was given 21 trials per day: seven choice-trials on which preference was measured, with both levers being presented and either response reinforced; and 14 guided trials with the two targets individually, which brought the total number of reinforcements to 14 for one of the targets and 7 for the other. For each animal, the majority stimulus was the one *against* which it had been trained in the previous experiment.

The results are shown in the first portion of Fig. 6. Each of the 14 animals developed a substantial preference for one of the two targets (8

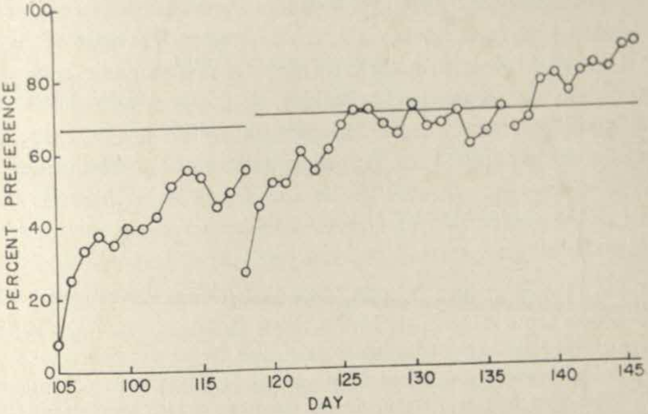


FIG. 6. PERFORMANCE IN EXPERIMENT IV

(In the first stage of training, all responses were reinforced, and guided trials were used to produce a 66:33 reinforcement-ratio. In the second stage, choices were reinforced on a 50:50 basis, and additional guided trials were used to produce a 70:30 reinforcement-ratio. In the third stage, choices were reinforced on a 100:0 basis, and guided trials were used to maintain the over-all 70:30 reinforcement-ratio.)

for the majority target and 6 for the minority target), the curve for the group as a whole therefore showing no preference. Under similar conditions, Ramond's rats developed a transient group-preference for the majority target which disappeared with further training, but not even a *transient* group-preference was shown by the fish. Whatever this difference may mean, the results for the fish indicate that controlling the distribution of reinforcements with guided trials is not enough to produce matching.

---

[10] C. K. Ramond, Performance in selective learning as a function of hunger, *J. exp. Psychol.*, 48, 1954, 265-270.

The second stage of this experiment was designed to test the hypothesis that guided reinforcements would be effective only if they directly followed unrewarded choices. All animals were given 20 trials per day. Ten were 50:50 choice-trials with guidance following unreinforced choices. The rest were randomly interspersed guided trials, one to the target preferred in the preceding phase of the experiment and nine to the other target. If the animals were influenced by *all* guidances, they should match at the 70% level, because there were, in all, 14 reinforcements to one target and 6 reinforcements to the other. If only the guidances which followed unreinforced choices were effective, the animals should match at the 50% level.

The results are plotted in the middle portion of Fig. 6. As the curve shows, matching at the 70% level appeared. Averaged over Days 126-136, the individual preferences for the more frequently reinforced stimulus range from 41-81% with a mean of 68%. (Without one animal which began with a very strong preference for the minority stimulus, the range of individual preferences is 61-81% and the mean is 70%. The preference of the deviant animal for the majority stimulus increased progressively from day to day to reach the 70% level on Day 136.) The results suggest that all of the guided reinforcements were effective, not only those which immediately followed unreinforced choices.

In the third stage of the experiment, each animal again was given 20 trials per day. Ten were 100:0 choice-trials with guidance after error, the majority stimulus of the previous stage being consistently reinforced. The rest were randomly interspersed guided trials so distributed as to maintain a 70:30 ratio of reinforcement; that is, there were four reinforcements for the target which was consistently reinforced on the choice-trials and six reinforcements for the other target.

The results are plotted in the third portion of Fig. 6. From the matching of the second stage of the experiment, there is an unmistakable shift to maximizing in the third, with every fish showing a marked and progressive increase in preference for the majority stimulus. The conclusion suggests itself that some inconsistency of reinforcement on choice-trials is essential if supplementary guided reinforcements are to influence choice.

Summary and Conclusions

In a series of experiments with a confounded visual-spatial discrimination, probability-learning in the African mouthbreeder was studied. When guidance was used to control the distribution of reinforcements between two stimuli—*i.e.* when, after each unreinforced response, the positive stimulus alone was introduced and the animal reinforced for response to it—probability-matching appeared in 80:20, 70:30, 60:40, and 50:50 problems. Precision of matching, as measured by the deviation of individual

preferences from group-values, increased progressively as the ratio of reinforcement decreased from 100:0 to 50:50. The results confirm and extend earlier findings on probability-learning in the mouthbreeder, lending emphasis to what seems to be an important functional difference between fish and mammal.

With guidance eliminated—*i.e.* when a simple noncorrectional method was employed—matching gave way to maximizing. Matching also disappeared when nonreinforcement was eliminated and the ratio of reinforcement maintained with guided trials alone—*i.e.* with reinforced responses to individually presented stimuli; nor could matching be demonstrated when a given stimulus was consistently positive on choice-trials and an intermediate ratio of reinforcement was maintained with interspersed guided trials. Some inconsistency of reinforcement apparently is necessary on choice-trials if the fish is to develop a preference-ratio corresponding to the over-all ratio of reinforcement. When 50:50 choice-trials were mixed with guided trials to give an over-all 70:30 ratio of reinforcement, a 70:30 distribution of choices developed. Clearly, the context in which reinforcement is given has considerable importance. Results of this sort highlight the limitations of the so-called mathematical theories which have been developed to deal with behavior in choice-situations.[11]

In two experiments, the ease of habit-reversal after 100:0 as compared with 70:30 training was studied, in one case when the use of guidance had led to different pre-reversal asymptotes, and in another case when the elimination of guidance had led to maximizing in both groups. In neither experiment was reversal retarded by inconsistent reinforcement. Directly analogous mammalian experiments are needed to assess the relation of these results to the data on extinction following partial reinforcement in simple instrumental situations.

---

[11] R. R. Bush and Frederick Mosteller, *Stochastic Models for Learning*, 1955, 1-365; R. R. Bush and W. K. Estes, (eds.), *Studies in Mathematical Learning Theory*, 1959, 1-432.

EFFECTS OF SLEEP-DEPRIVATION AND CHLORPROMAZINE ON SIZE-CONSTANCY JUDGMENTS

By V. R. CARLSON, National Institute of Mental Health

In a previous study, the hypothesis was developed that overestimation in the size-constancy experiment is due to *O*'s assumption of the perspective attitude and his motivational reaction to the experimental situation.[1] Briefly, the argument is that instructions to make a match in terms of objective size favor a bias leading to overestimation of the size of a distant object. Instructions to make a match on the basis of apparent visual size tend to eliminate this bias, provided that the implication of an analytic task is not introduced, resulting on the average in an accurate physical match with no systematic error as a function of distance. This difference between instructions of apparent- vs. objective-size was demonstrated in the previous experiment. The present study explores the motivational aspect of this formulation.

People deprived of sleep can and often do perform at or near normal levels of competence, although they may not be able to maintain an efficient level of continuous performance for very long.[2] This seeming lack of impairment in capacity often has been attributed to increased effort which compensates for losses that would otherwise be evident. Without implying that increased effort actually does operate to better performance, it seems quite likely that *O*'s expectancy is one of impaired functioning for which he must compensate if he is to perform as well as normal. Given also the assumption that the natural criterion for competent performance in the size-constancy situation is the ability to make an accurate judgment of physical size, then sleep-deprivation, as a motivating condition, should favor a response-set similar to that normally introduced by objective in-

* Received for publication July 2, 1960. The author is indebted to Mr. Eugene Tassone for assistance in the collection and analysis of the data.

[1] V. R. Carlson, Overestimation in size-constancy judgments, this JOURNAL, 73, 1960, 199-213.

[2] A. S. Edwards, Effects of the loss of one hundred hours of sleep, this JOURNAL, 54, 1941, 80-91; Nathaniel Kleitman, *Sleep and Wakefulness*, 1939, 300-321; D. B. Tyler, The effect of amphetamine sulfate and some barbiturates on the fatigue produced by prolonged wakefulness, *Amer. J. Physiol.*, 150, 1947, 261; R. T. Wilkinson, *Effects of Lack of Sleep*, Medical Research Council, Cambridge, England, Flying Personnel Research Committee Rep. No. 961.3, 1956, 1-3.

structions. According to the present argument, the result would be overestimation of actual size in spite of instructions to match for apparent size.

Williams, Lubin, and Goodnow have been able to reconcile many of the apparently inconsistent results of studies on sleep-deprivation on the basis of an hypothesis of lapse.[3] Rather than a gradual and continuous impairment in performance, intermittent periods of lowered responsiveness are postulated. Qualitatively normal performance can be achieved between such lapses, and, if the task does not require response during lapses there may be no evidence of impairment. In this sense, the task of judging size-constancy avoids the probability of a decrement in performance, since judgments are made only when *O* is relatively alert and responding.

The psychological state induced by chlorpromazine is similar in a number of respects to that resulting from sleep-deprivation. With chlorpromazine, however, *O* exhibits little, if any, evidence of a compensatory drive to make up for possible deleterious effects of the drug on performance.[4] Chlorpromazine would, therefore, appear to be an appropriate comparison for the postulated motivational effect of sleep deprivation on judgments of size.

METHOD

This experiment was undertaken in conjunction with an evaluation of the effects of dextro-amphetamine on some other kinds of performance with the same *O*s.[5] Although no hypothesis with respect to possible effects of dextro-amphetamine on judgments of size-constancy was being tested, the judgments of size were also obtained under that condition, and they always followed approximately after an hour of testing on several other visual-motor tasks. No discernible complications arose, however, due to the combining of the two experiments.

The task, apparatus, and stimulus-conditions were the same as those used in an earlier study.[6] Administration of drugs and placebo was oral, in identical capsules, and the *O*s did not know what particular drugs were being used.

Observers. The *O*s were 7 men and 7 women, 18–25 yr. of age, from a population of normal volunteers for medical research. *O*s from this population have been compared with college students on the procedures used here and no significant differences in performance were found. One *O*, a woman, exhibited undue elevation in blood-pressure during the first sleep-deprivation session, and it was deemed

[3] H. L. Williams, Ardie Lubin, and J. J. Goodnow, Impaired performance with acute sleep loss, *Psychol. Monogr.*, 73, 1959, (No. 484), 22-24.

[4] J. A. Starkweather, Chlorpromazine, reserpine, and meprobamate: psychological effects, in J. O. Cole and R. W. Gerard (eds.), *Psychopharmacology, Problems in Evaluation*, 1959, 501-506; Abraham Wikler, *The Relation of Psychiatry to Pharmacology*, 1957, 27-38.

[5] Conan Kornetsky, A. F. Mirsky, E. K. Kessler, and J. E. Dorff, The effects of dextro-amphetamine on behavioral deficits produced by sleep loss in humans, *J. Pharmacol.*, 127, 1959, 46-50.

[6] Carlson, *op. cit.*, 201-203.

medically inadvisable for her to continue in the experiment. She did not, therefore, participate in the second sleep-deprivation and chlorpromazine conditions.

Task and apparatus. *O* adjusted a variable triangle at a distance of 10 ft. to match a standard triangle at 40 ft. The lateral separation between the two was 20°, and both were approximately at *O*'s eye-level. The task was performed under full illumination with free binocular regard in a 10 × 50-ft. room containing numerous miscellaneous objects.

Instructions. For the judgments of apparent size, *O* was told:

> So adjust the variable triangle that it looks equal to the standard in apparent visual size. It may also be equal in actual physical size at that point, or it may not —we are not concerned about that. Try to adjust the variable that it appears equal to you visually, whether you think it is equal in actual size or not.

For objective size, *O* was told:

> So adjust the variable triangle that it is, as best you can judge, equal in actual physical size to the standard,—that, if you were to measure them both with a ruler, they would be the same size. They may also look equal to you in apparent visual size or they may not—we are not concerned about that. Try to adjust the variable that you think it is the same actual size as the standard whether it appears equal to you visually or not.

These instructions were developed to satisfy the following criteria: (a) if there is no natural tendency for *O* to differentiate conceptually between apparent size and physical size, he must not be forced to do so artificially, and (b) if *O* does make a distinction between these two concepts, he must not be left in doubt as to which one he should apply in making his response.

After the first day *O* was also asked: "How does the (standard) triangle look today? Can you see it clearly? Does it look any different to you today, or the same as before?" No special effort was made, however, to elicit subjective reports beyond the answers *O* volunteered to these questions.

The triangles, isoceles with an altitude-to-base ratio of 1.39, were cut from white cardboard. Each was mounted on an apparatus which allowed the triangle to be adjusted vertically through a slot against a 30 × 36-in. black felt background. The standard triangle was set at three primary values, 44, 117, and 159 mm. in altitude. Four additional values were used on both sides of 117 mm., spaced in steps of 3.2 mm. These were included to maintain the appearance of a relatively large number of standards and also to obtain a measure of sensitivity to small differences in the size of the standard. The 11 values of the standard were presented twice in randomized order at each session.

Sleep-deprivation. Maintenance of wakefulness began at 8:00 A.M. on Sunday of a given week and continued until 8:00 A.M. the following Wednesday. The testing sessions were begun 45.5 and 69.5 hr. after the start of sleep-deprivation, the size-constancy judgments were obtained at approximately 47 and 71 hr. The same schedule was repeated with the same group of *O*s during the next Sunday-to-Wednesday period. The *O*s were sleep-deprived in groups of 3–4 on succeeding two-week periods. During sleep-deprivation, the *O*s were under continuous observation and supervision by the nursing and attendant staff on a ward equipped and operated similarly to that of a modern hospital. The *O*s had free access to the recreation and day-rooms, and were allowed to engage in activities such as painting, playing cards or pool, and so forth.

Chronological order of conditions. The order of conditions was as follows:

(A) Apparent-size instructions:
  (1) Placebo;
  (2) Sleep-deprivation;
    (a) Placebo at 44 hr.;
    (b) Placebo at 68 hr.;
  (3) Sleep-deprivation;
    (a) Dextro-amphetamine (10 mg.) at 44 hr.;
    (b) Dextro-amphetamine (15 mg.) at 68 hr.;
  (4) Placebo;
  (5) Chlorpromazine.
(B) Objective instructions:
  (1) Control (no drug, placebo, nor sleep-deprivation);
  (2) Control (no drug, placebo, nor sleep-deprivation).

Several days to a week elapsed between successively numbered conditions.

Secondary control experiments. Chlorpromazine constricts the pupil without producing any noticeable impairment of vision. *O*s under sleep-deprivation sometimes have reported visual symptoms such as blurring or diplopia, but impairment on tests of purely visual function is apparently not a significant accompaniment of sleep-loss.[7] Even if some oculomotor impairment did occur, it would probably not have important effects on judgments of size under the present conditions, since the images of the standard and variable test-objects would not only have to be affected differentially with respect to each other but probably also differentially with respect to the immediate context of visual stimulation in which each occurs. Furthermore, it is diffiult to deduce a predictable consequence of visual impairment on the size-judgments except on the basis of a general reduction in the normal cues to depth-perception. Reduction-conditions favor matches tending in the direction of projective size, a tendency which would be opposite to the effect hypothesized here. Nevertheless, this matter was checked with two additional groups of *O*s from the same population as those in the main experiment. These *O*s performed the same size-constancy task with apparent-size instructions after topical administration of drugs to the eyes.

Ten *O*s received three successive instillations of two drops in each eye of the following drugs on different days: pilocarpine (1% solution), Neo-Synephrine (10% solution), and sodium chloride (0.9% solution). Another group of 16 *O*s was similarly administered homatropine (2% solution) on one day and sodium chloride (0.9% solution) on another day. Pilocarpine constricts the pupil and causes spasm of accommodation. Homatropine dilates the pupil and causes paralysis of accommodation. Neo-Synephrine dilates the pupil with essentially no effect on accommodation.[8] Administration of saline served as a control condition.

The order of receiving these drug-conditions was counterbalanced among the *O*s, and at least one day intervened between successive conditions. At each experimental session, pupil-size was measured by comparing *O*'s pupil with a series of

---

[7] Edwards, *op. cit.*, 88; Kleitman, *op. cit.*, 305-306; Tyler, *op. cit.*, 254-255; Wilkinson, op. cit., 1.

[8] The effects of these drugs on ocular-motor adjustment are discussed in detail by K. W. Christoferson and K. N. Ogle, The effect of homatropine on the accommodation-convergence association, *A. M. A. Arch. Ophthal.*, 55, 1956, 779-791; F. C. Sabin and K. N. Ogle, Accommodation-convergence association, *A. M. A. Arch. Ophthal.*, 59, 1958, 324-332.

black solid circles drawn on a card held next to the eye, and visual acuity was tested with a Bausch and Lomb Ortho-rater.

Results

A score of size-constancy for each *O* under each condition was computed as the average ratio of *O*'s setting of the near variable to the actual size of the far standard for the three main sizes of the standard triangle (44, 117, and 159 mm.). A ratio greater than 1.00 indicates overestimation of the standard. Reliability coefficients for the various conditions (correlations between the scores obtained on different days) ranged from 0.82 to 0.95. There were no significant mean-differences between the first

TABLE I

Average Ratios of *O*s' Settings of Near Variable to Actual Sizes of Far Standard

($N=14$, except for chlorpromazine where $N=13$.)

Condition	Mean	*SD*	Significant differences*				
			1.00†	(1)	(2)	(3)	(4)
Apparent-size instructions							
(1) Placebo	1.12	0.16	+	−	+	−	+
(2) Sleep-deprivation	1.27	0.20	+	+	−	+	−
(3) Chlorpromazine	1.09	0.23	−	−	+	−	+
Objective-size instructions							
(4) Control	1.32	0.24	+	+	−	+	−

* The *t*-test between correlated means was used and $p<0.05$ was taken as the confidence level. Significant values are not altered if data for the *O* not present in the chlorpromazine-condition are omitted from all comparisons.

† Value for objectively correct match.

and second placebo-conditions, between the two control sessions with objective instructions, nor among the four conditions of sleep-deprivation. The separate values within each of these categories have therefore been averaged (Table I).

Placebos produced a statistically significant degree of overestimation, but sleep-deprivation produced greater overestimation, comparable to that obtained with objective instructions. The results with chlorpromazine were not significantly different from those with placebos or from a ratio of 1.00.

The slightly differing values of the standard around 117 mm. were utilized to obtain two measures of precision within-*O*s, computed for each *O* separately and then averaged across *O*s; viz. (a) the *SD* of the ratios of *O*'s settings to the values of the standard; and (b) the rank-order correlation between *O*'s settings and the values of the standard. These measures did not exhibit satisfactory reliability (average $r_{II} = 0.46$ and

0.32 for the variability and correlational measures, respectively). They are the only scores in the present data, however, which reflect any effect which can unequivocally be termed a decrement or impairment in performance. The several sleep-deprivation conditions were, again, not significantly different from each other, but the average sleep-deprivation scores were significantly different from those obtained with placebos (Table II). Under sleep-deprivation, the *O*s were somewhat more variable and less sensitive to small differences in the standard triangle. Chlorpromazine and objective instructions were without significant effect on these measures.

There was ample confirmation of the presumption that the *O*s would meet the sleep-deprivation condition as a challenge. All *O*s displayed a good deal of initiative in engaging in activities which would help keep them awake, with very few attempts at circumventing the no-sleep requirement, and no instances of overt reluctance to continue in the project. During a series of size-constancy trials, *O* often would state that he was becoming drowsy and ask permission to stand up and walk around. Most *O*s failed to respond to the go-ahead signal on some trials. In these instances, *O* was easily aroused, and he frequently voiced determination to try harder.

TABLE II

WITHIN-*O*S PRECISION AND SENSITIVITY TO SMALL CHANGES IN THE STANDARD TRIANGLE ($N=14$)

	Variability*		Correlation†	
	Placebo	Sleep-deprivation	Placebo	Sleep-deprivation
Mean	0.084	0.101	0.707	0.539
SD	0.023	0.023	0.172	0.173
p‡	<0.001		<0.01	

* *SD* of the ratio of *O*'s setting to the value of the standard for the 9 values varying from 104.2 to 129.8 mm.

† The coefficients of the rank-order correlation between *O*'s settings and the nine values of the standard varying from 104.2 to 129.8 mm.

‡ Probability associated with the *t*-value for the difference between correlated means.

The contrast with chlorpromazine was great. With this drug, all *O*s were passively resistant toward performing the task and expressed a desire to go to bed and be left alone. Nearly all had to be urged into doing the task, and some refused at first, claiming inability to walk to the experimental room, whereupon they would be taken there in wheel-chairs.[9] Once the trials were actually started, all *O*s appeared

[9] A common symptom of large doses of chlorpromazine is muscular weakness, and the dosage used (200 mg.) should be considered large. Total daily doses as high as 1600 mg., however, have been used with some patients (L. S. Goodman and A. Gilman, *The Pharmacological Basis of Therapeutics*, 1955, 1067).

capable of executing the task, and their complaints were directed toward the task being "trying" and seeming to "drag on."

The general response to the questioning of *O* as to the visual appearance of the standard triangle during the sleep-deprivation and chlorpromazine conditions was that it looked about the same as in the first placebo-condition. With sleep-deprivation, some *O*s reported that greater effort was required to keep the triangle single-imaged and in focus, others did not. At their last sleep-deprivation session, two *O*s reported that the standard triangle looked farther away than it had previously, but they made the same comment at the following (non-sleep-deprivation) placebo-session. Only one *O* reported an hallucination (a dancing dwarf which appeared briefly several times). There was no discernible relationship between subjective reports and *shift* in size-constancy score from the placebo-value. The *O*s who reported subjective visual symptoms were, however, those who tended to produce the higher size-constancy ratios in all conditions.

The control-experiments concerned with possible effects of changed ocular-motor function were completely negative. Although clear drug-effects on pupil-diameter and acuity-scores were observed, there were no significant differences in the judgments of size-constancy. The average size-ratio for all these conditions was 1.02, which was not significantly different from a value of 1.00.

Discussion

The high motivational state produced in this study by the experimental deprivation of sleep is descriptively similar to that reported by Morris, Williams, and Lubin.[10] Our *O*s yielded, however, fewer subjective symptoms than theirs. This might be due either to differences in sampling or to the fact we did not urge our *O*s to elaborate upon the subjective aspects.

The moderate overestimation of size made in the control (placebo) condition may represent a placebo-effect, since significant practice- and order-effects were not found in the previous experiment under the same stimulus-conditions.[11] Such placebo-effect would not be inconsistent with the present motivational hypothesis; but this overestimation might also be a matter of error in the sampling of *O*s. An appropriate control by means of which it could be evaluated was not included. Otherwise, the theoretical effects of sleep-deprivation and chlorpromazine seem to have been obtained.

If *O* is more highly motivated to perform well, he tends, apparently, to overestimate the size of a distant object relative to a nearby object when he is asked explicitly to judge size. This generalization appears to

---

[10] G. O. Morris, H. L. Williams, and Ardie Lubin, Misperception and disorientation during sleep deprivation, *A. M. A. Arch. gen. Psychiat.*, 2, 1960, 248.

[11] Carlson, *op. cit.*, 204.

apply also to the results of Singer, who found that experimental frustration changed his *O*'s judgments of size-constancy in the direction of overestimation.[12] The circumstances of both experiments qualitatively suggest a defensive or compensatory reaction rather than a simple increase in effort, an interpretation which may also be consistent with the results of Sanders and Pacht.[13] The latter investigators found a positive relation between degree of psychopathology and amount of overestimation. This does not mean that all psychotic patients will overestimate. In fact, they do not.[14] It is interesting, however, that patients most likely to exhibit a superficially well-integrated but defensively compensative response—paranoid individuals—have shown overestimation in size-constancy judgments.[15]

It is important to realize that these experimental results do not imply motivational, personological, or attitudinal effects on perceptual constancy as it occurs in functional commerce with objects in the individual's habitual, everyday environment. The present rationale assumes just the opposite; namely, that perception under naturally-occurring circumstances is very generally veridical. On the basis of the perspective attitude, a nearer object should subjectively 'look' larger than a more distant one of equal size. If *O* were in a neutral motivational state with respect to the quality of his performance in the experimental situation, the two equal-sized objects would satisfy the perspective attitude, for the more distant one can be judged to occupy a smaller proportion of the visual field,[16] and no corrective allowance for distance would be necessary in either an objective or a subjective sense.

The present results with chlorpromazine would seem to be a particularly appropriate illustration of this argument. It is difficult to imagine any more neutral motivational state with respect to the experimental task than was exhibited by these *O*s. Under such circumstances it is postulated that *O*'s response is most purely a function of the immediate, proximal stimulus-situation, uninfluenced by attitudinal response-biases. This indifferent psychological state is what, in the writer's opinion, *E* is usually trying to

---

[12] J. L. Singer, Personal and environmental determinants of perception in a size constancy experiment, *J. exp. Psychol.*, 43, 1952, 420-427.

[13] Richard Sanders and Asher Pacht, Perceptual size constancy of known clinical groups, *J. consult. Psychol.*, 16, 1952, 440-444.

[14] Edward Lovinger, Perceptual contact with reality in schizophrenia, *J. abn. soc. Psychol.*, 52, 1956, 87-91; T. E. Weckowicz, Size constancy in schizophrenic patients, *J. ment. Sci.*, 103, 1957, 475-486.

[15] H. L. Raush, Perceptual constancy in schizophrenia: I. Size constancy, *J. Personal.*, 21, 1952, 176-187.

[16] R. B. Joynson, The problem of size and distance, *Quart. J. exp. Psychol.*, 1, 1949, 119-135.

induce when he exhorts *O* to respond on the basis of how the stimulus-objects 'look,' a directive which is communicatively ambiguous at best.

The fact that size-judgments appear to approximate constancy most closely in a neutral motivational state is of fundamental importance for perceptual theory. It suggests that perception of object-size does not start from angular retinal-image size, as such, and end up as a compromise between angular size and physical object-size.[17] Rather, veridical object-perception is the basic starting point, and judgments experimentally obtained can be influenced in either direction. If the influence, instructional or situational, is toward objective accuracy, 'overconstancy' is the result. If the influence is toward a judgment of how things 'look,' as distinguished from physical reality, then 'underconstancy' occurs.

Both deviations can be derived by assuming the operation of the perspective attitude, which says that a more distant object appears smaller than a nearer one of equal physical size. To achieve an objective match, there is no precise limit as to how large the nearer object can be, but it clearly cannot appear to be smaller in physical size without violating the perspective attitude. There is a one-sided opportunity for bias in the direction of 'overconstancy.' To achieve a match in terms of 'apparent' size, implying something different from objective size, there is no clear limit as to how small the nearer object can be relative to the farther one, but it cannot appear to be larger in physical size. There is therefore a one-sided opportunity for bias in the opposite direction.

SUMMARY

On the hypothesis that overestimation in judgments of size-constancy depends upon *O*'s assumption of the perspective attitude and his motivational effort toward objective accuracy, it was predicted that sleep-deprivation would produce overestimation of size at a distance. Chlorpromazine was utilized as a low-motivation, condition of comparison. Fourteen normal *O*s made size matches after 47 and 71 hr. of sleep-deprivation, and following administration of placebos and 200 mg. of chlorpromazine. Some overestimation of size occurred with placebos, but sleep-deprivation produced a significantly greater degree of overestimation. The effect of chlorpromazine was less than, though not significantly different from, that of the placebos. Additional evidence was obtained indicating that the results were not due to possible effects of the experimental conditions on oculomotor functioning.

[17] C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 276.

ALTERNATING AND RANDOM PARTIAL REINFORCEMENT IN THE FISH, WITH SOME OBSERVATIONS ON ASYMPTOTIC RESISTANCE TO EXTINCTION

By R. C. Gonzalez, Rochelle M. Eskin, and M. E. Bitterman, Bryn Mawr College

In a number of recent experiments with the fish *Tilapia macrocephala* (the African mouthbreeder), initial resistance to extinction was found to be greater after consistent than after partial reinforcement.[1] Only in the course of a series of extinctions separated by reconditioning sessions did evidence of the paradoxical effect which is characteristic of the rat make an appearance: resistance to extinction declined progressively both in partially and in consistently reinforced fish, but more rapidly in the latter, and their resistance at asymptote was less than that of the partially reinforced animals.[2] The research which is reported here extends the comparison of fish and rat in terms of the consequences of partial reinforcement.

Resistance to extinction in the rat is greater after random 50% reinforcement than after alternating 50% reinforcement, which suggests that the rat is capable of discriminating patterns of reinforcement, and this interpretation is supported by the fact that resistance to extinction actually may decrease as the amount of training with alternating reinforcement increases.[3] A trial-by-trial analysis, such as that of Sheffield and Hull, leads, of course, to the expectation of precisely the opposite results.[4] One pur-

* Received for publication June 29, 1960. This research was supported by Grant M-2857 from the National Institute of Mental Health. Dr. Gonzalez participated in the work during his tenure of a Postdoctoral Fellowship awarded by the same Institute.

[1] Jerome Wodinsky and M. E. Bitterman, Partial reinforcement in the fish, this Journal, 72, 1959, 184-199; Nicholas Longo and M. E. Bitterman, The effect of partial reinforcement with spaced practice on resistance to extinction in the fish, *J. comp. physiol. Psychol.*, 53, 1960, 169-172; Wodinsky and Bitterman, Resistance to extinction in the fish after extensive training with partial reinforcement, this Journal, 73, 1960, 429-434.

[2] Wodinsky and Bitterman, *op. cit.*, this Journal, 72, 1959, 184-199.

[3] D. W. Tyler, E. C. Wortz, and M. E. Bitterman, The effect of random and alternating partial reinforcement on resistance to extinction in the rat, this Journal, 66, 1953, 57-65; E. J. Capaldi, The effect of different amounts of training on the resistance to extinction of different patterns of partially reinforced responses, *J. comp. physiol. Psychol.*, 51, 1958, 367-371.

[4] V. F. Sheffield, Extinction as a function of partial reinforcement and distribution of practice, *J. exp. Psychol.*, 39, 1949, 511-526; C. L. Hull, *A Behavior System*, 1952, 120-121.

pose of the present research was to determine whether the fish is affected in like manner by the pattern of partial reinforcement.

Another purpose of this research was to examine the influence of interpolated reconditioning on asymptotic resistance to extinction. When rat or fish is subjected to a series of extinction-sessions separated by reconditioning sessions, resistance to extinction falls progressively to a low level, and, in the rat, that level may be independent of the amount of interpolated reconditioning. After a series of regularly alternated conditioning and extinction sessions, Wickens and Miles randomized the order of the two types of session without affecting resistance to extinction; that is, resistance was no greater after a reconditioning session than after an extinction-session.[5] A similar experiment with the fish did not, however, yield comparable results,[6] and some observations made at the conclusion of the first phase of the present work led to a further study of the problem.

METHOD

Subjects. The *Ss* were 20 experimentally naïve African mouthbreeders of both sexes, about 2.5 in. long and about 6 mo. old. The animals were bred in group tanks, and, after reaching sexual maturity, they were maintained in individual 2-gal. tanks on open shelves in the laboratory. Studied also were 11 experimentally naïve, male rats of the California maze-bright strain, about 6 mo. old. They were bred in the laboratory and housed in individual cages.

Apparatus. The apparatus employed for the fish has been described in detail elsewhere.[7] It consisted of a target which could be lowered into *S*'s individual living tank, a system for detecting *S*'s contacts with the target, and a pellet-feeder operated by the detecting system. A Standard Electric timer was used to measure the latency of response when discrete trials were given; the clock started with the introduction of the lever and was stopped by the response. In work with a continuously available lever, a counter was used to measure the number of responses which occurred in a given time.

The apparatus used for the rat was a Skinner box constructed of plastic. The animal's compartment was 9 in. long, 7 in. wide, and 7 in. high. The floor was made of brass rods, although no shock was used in this experiment. The lever was a stirrup-shaped rod inserted through the roof of the compartment to a point 4 in. above the floor; the height of the bar increased the difficulty of pretraining to some extent but markedly reduced the number of responses registered in the course of general activity. The pellet-feeder was of the same design as that used for the fish.[8] The number of responses made during a given period of time was read from a counter.

Procedure. (A) After being trained to strike the target, all fish were given 20

[5] D. D. Wickens and R. C. Miles, Extinction changes during a series of reinforcement-extinction sessions, *J. comp. physiol. Psychol.*, 47, 1954, 315-317.
[6] Wodinsky and Bitterman, *op. cit.*, this JOURNAL, 72, 1959, 184-199.
[7] Longo and Bitterman, Improved apparatus for the study of learning in fish, this JOURNAL, 72, 1959, 616-620.
[8] *Ibid.*, 619.

daily sessions of 20 discrete trials each, with all responses reinforced, and an inter-trial interval of about 2 sec. Then the animals were divided into two equated groups and given 20 more days of training which differed from the training of the first 20 days only in that partial reinforcement was introduced. For the Alternating Group, odd trials were reinforced and even trials were unreinforced. For the Random Group, the pattern of reinforced and unreinforced trials followed Gellermann-orders, with the restriction that the first trial of each day was reinforced and the last trial unreinforced.[9] On Day 41, all *S*s were extinguished to the criterion of five successive failures to respond in 30 sec. When, on any trial, the animal failed to respond in that time, the target was withdrawn, the time was recorded, and the target was re-inserted as usual to begin the next trial. After this first extinction, there were three reconditioning sessions (with the same patterns of reinforcement as before), a second extinction to the same criterion, three more conditioning sessions, and so on, until the animal had been extinguished seven times. Throughout the experiment, the fish were maintained on a schedule of 20 pellets per day, and any portion of this ration which was not earned in the experimental situation was given 1 hr. later in the living area.

(B) Two new groups (Group I and Group II) then were constituted, using matched members of the original Alternating and Random Groups. There were eight animals in each group, several having been lost earlier in the experiment. The treatment of the animals of Group II continued as before for four more extinctions, each preceded by three days of conditioning. For the animals of Group I, however, all conditioning trials now were reinforced.

(C) All animals were given three days of training with a continuously available target (the Skinnerian method). All responses were reinforced, and the time taken to earn the 20 reinforcements available on each day was measured. In the fourth Skinnerian session, there was no reinforcement, and the number of responses in 10 min. was measured. Then there were three more days of continuous reinforcement, another 10-min. extinction-session, and so forth, until there had been six Skinnerian extinction-sessions in all. The purpose of this part of the experiment was to prepare the animals for the next part, in which resistance to extinction as a function of the amount of interpolated reconditioning was measured.

(D) In this part of the experiment, only 12 of the original fish were used, since the design called for a number of *S*s which was a multiple of 6. Each *S* was given three extinctions, one after 0, one after 3, and one after 6 days of conditioning, in a counterbalanced order. During conditioning, reinforcement was continuous, and the time taken to earn 20 reinforcements was measured. Upon the completion of the first cycle of testing, which lasted 12 days, a second cycle was conducted in the same manner.

For purposes of comparison, a similar experiment was done with rats. The rats were reduced to 80% of their satiated bodily weights and maintained on a 24-hr. feeding schedule. After a period of adjustment to the apparatus, and experience with a baited lever, all *S*s were given 15 days of continuous reinforcement, on each of which the time to earn 30 pellets was measured. On the next day, there was an extinction-session, and the number of responses made in 10 min. was measured. There followed three more days of conditioning, another extinction, and so forth,

[9] L. W. Gellermann, Chance orders of alternating stimuli in visual discrimination experiments, *J. genet. Psychol.*, 42, 1933, 206-208.

until the animals had been subjected to seven extinction-sessions of the same kind. After the seventh extinction, each rat was given three additional extinctions, one after 0, one after 3, and one after 6 days of conditioning, in approximately balanced fashion; since there were only 11 rats, 2 were assigned to 5 of the 6 possible orders and only a single animal to the sixth order (6-0-3).

RESULTS

(A) With consistent reinforcement, latency of response in the *fish* declined in negatively accelerated fashion to an asymptote of about 0.23 sec. The introduction of partial reinforcement had no significant effect on mean latency, although *both* groups developed a strong tendency to slow

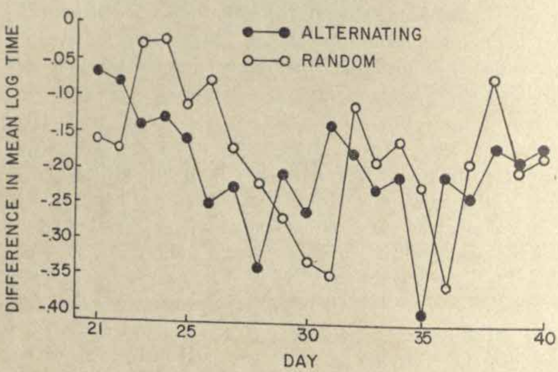


FIG. 1. DIFFERENTIAL RESPONSE AFTER REINFORCEMENT AND NONREINFORCEMENT IN ALTERNATELY AND RANDOMLY REINFORCED FISH

(A negative difference-score reflects less rapid response after reinforcement than after nonreinforcement.)

down following reinforced trials and to speed up following unreinforced trials. The difference in mean log time for trials following reinforcement and nonreinforcement on each day of training is plotted in Fig. 1. The two curves seem very much the same, and this impression is substantiated by an analysis of variance (Lindquist's Type I) which shows no significant difference between the two groups ($F < 1$ with 1 and 14 *df.*).[10] In extinction, too, the performance of the Alternating and Random *Ss* was very much the same. Plotted in the first portion of Fig. 2 is mean number of responses to criterion for the two groups in each of the first seven extinctions. Although a Type-I analysis shows a significant decline in resistance over the series of extinctions ($F = 10.22$ with 6 and 84 *df.*, $p <$

[10] The analysis is based on the data for 8 *Ss* in each group, 4 *Ss* having been lost in the course of training.

0.01), the difference between groups is negligible ($F < 1$ with 1 and 14 *df.*).

These results provide an interesting contrast with those for the rat. Alternating and random 50% partial reinforcement have different effects on the rat. A tendency to respond more rapidly after nonreinforcement than after reinforcement develops in training with alternating but not with random reinforcement, and resistance to extinction is less after alternating than after random reinforcement. In the fish, the two patterns of reinforcement do not have differential effects. A tendency to respond more rapidly after nonreinforcement than after reinforcement—which perhaps may be interpreted as a frustration-phenomenon—appears under both conditions, and resistance to extinction following the two patterns of reinforcement is the same. In situations such as the one used here, apparently, it is the number of reinforcements rather than the pattern of reinforcement which is important for the fish.

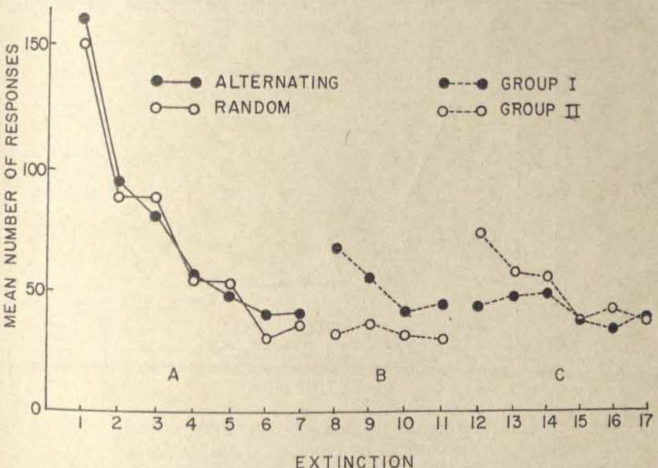


FIG. 2. RESISTANCE TO EXTINCTION IN PHASES *A, B,* AND *C* OF THE WORK WITH FISH

(B) In the second phase of the experiment, consistent reinforcement was introduced for half the animals and its effect on resistance to extinction was studied. The treatment of the newly constituted Group II (consisting of four Alternating and four Random fish) continued as before—three days of partial reinforcement followed by extinction, and so forth—and its resistance to extinction remained the same, as the central portion of Fig. 2 shows. The newly constituted Group I (consisting of four Alternating and four Random fish matched with those of Group II) was shifted to

consistent reinforcement, and its resistance to extinction rose substantially, at least to begin with. The mean difference between number of responses in the seventh extinction and the number of responses in the eighth extinction was significantly greater for Group I ($t = 2.35$, $p < 0.05$). Here is another indication of the tendency for resistance to extinction in the fish to be greater after consistent than after partial reinforcement.

These results bear also on the nature of the processes which determine the asymptote reached by the curve of resistance to extinction in a conditioning-extinction series. It has been assumed that the asymptote is de-

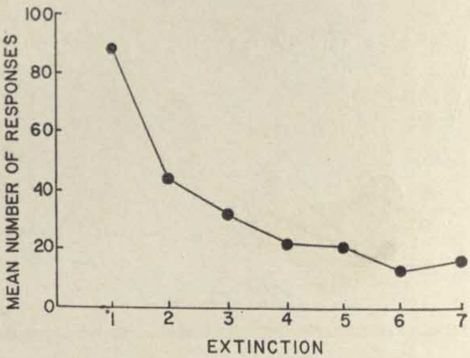


FIG. 3. CHANGES IN THE RAT'S RESISTANCE TO EXTINCTION IN THE COURSE OF A CONDITIONING-EXTINCTION SERIES

termined simply by the discriminability of conditioning and extinction days,[11] but the present results suggest that, at least in the fish, it is influenced as well by the amount of interpolated reinforcement. It would be unreasonable, certainly, to contend—as one would have to if one wished to rely alone on the discrimination-hypothesis—that the shift from partial to consistent reinforcement somehow made the discrimination between conditioning and extinction more *difficult* for the *S*s of Group I.

(C) The same dubious assumption would be required to account in terms of the discrimination-hypothesis for the results of the third phase of the experiment, in which both groups were consistently reinforced. The data are plotted in the right-hand portion of Fig. 2. The resistance to extinction of Group II, shifted from partial to continuous reinforcement, was greater at the outset than that of Group I. (By the Mann-Whitney test, which was indicated by skewness and by heterogeneity of variance, the dif-

[11] O. H. Mowrer and Helen Jones, Habit strength as a function of the pattern of reinforcement, *J. exp. Psychol.*, 35, 1945, 293-311; C. C. Perkins, Jr., and A. J. Cacioppo, The effect of intermittent reinforcement on the change in extinction rate following successive reconditioning, *J. exp. Psychol.*, 40, 1950, 794-801.

ference between the two groups in number of responses for the twelfth extinction is significant beyond the 5% level.) Some sort of transient contrast-effect seems required to account for these data.

(D) The results already presented suggest that it is difficult, even in a long conditioning-extinction series, to bring the fish to a point at which its resistance to extinction is unaffected by the amount of interpolated reinforcement. The phenomenon was studied further in the final phase of the experiment, and some additional data on the rat were collected as well.

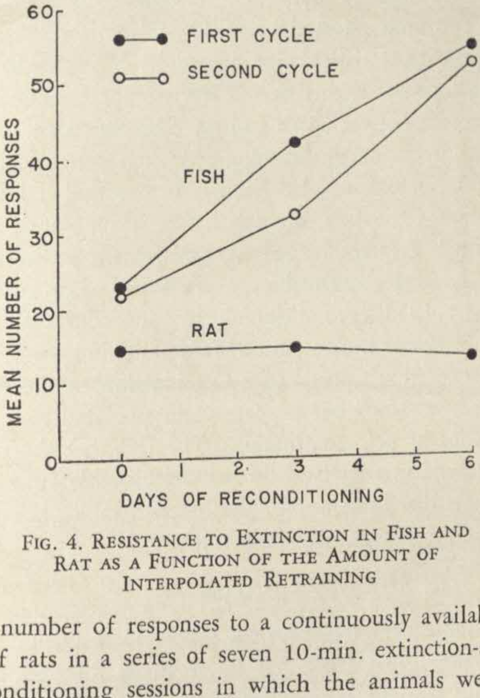


FIG. 4. RESISTANCE TO EXTINCTION IN FISH AND RAT AS A FUNCTION OF THE AMOUNT OF INTERPOLATED RETRAINING

The mean number of responses to a continuously available lever made by a group of rats in a series of seven 10-min. extinction-sessions (separated by reconditioning sessions in which the animals were consistently reinforced) is shown in Fig. 3. As in previous experiments with the rat, the curve falls sharply to a stable level.[12] In Fig. 4, responses during extinction are plotted both for fish and for rat as a function of the number of days of continuously reinforced retraining which antedated each extinction-session in the final phase of the experiment. The curve for the rat is perfectly flat, but the curves for the fish (based on an original set of measurements and a replication) show a marked positive relationship between number of responses in extinction and number of days of reconditioning.

[12] Perkins and Cacioppo, *op. cit.*, 794-801; D. H. Bullock and W. C. Smith, An effect of repeated conditioning-extinction upon operant strength, *J. exp. Psychol.*, 46, 1953, 349-352; Wickens and Miles, *op. cit.*, 315-317.

An analysis of Lindquist's Type II based on the data obtained in the first cycle of measurement shows the variance due to number of days of reconditioning to be significant beyond the 5% level ($F = 8.63$ with 2 and 6 *df.*). Neither order ($F < 1$ with 2 and 6 *df.*) nor the interaction between order and days of reconditioning ($F = 4.78$ with 2 and 6 *df.*) is significant, probably because there were only 2 *S*s in each order. A simple treatments × subjects analysis of the data for the second cycle shows the effect of number of days of reconditioning to be significant beyond the 1% level ($F = 10.12$ with 2 and 11 *df.*). The difference between the three-day values for the two cycles, tested by Wilcoxon's method for paired replicates, fails to approach statistical significance.

These results, together with those of earlier experiments, reflect another interesting difference between fish and rat. The resistance to extinction of the rat falls rapidly in a conditioning-extinction series to a point which seems to be independent of the amount of interpolated reinforcement. Comparable data for the fish suggest that such a point is reached much more slowly, if it is ever reached at all. Resistance to extinction in the fish falls in negatively accelerated fashion over a series of extinctions when the amount of interpolated reconditioning remains constant, but the level achieved is a function of the amount of reconditioning. In the fourth phase of the present experiment, resistance to extinction in the fish was influenced markedly by the number of days of interpolated reconditioning despite a long series of prior extinctions. The fact that the results for the second cycle of tests was statistically indistinguishable from those for the first cycle suggests that the slope of the function might not approach zero even if the testing were extended indefinitely.

SUMMARY

African mouthbreeders trained in a simple instrumental situation and subsequently extinguished were not differentially affected by alternating as compared with random 50% reinforcement. Both patterns of reinforcement yielded a tendency to respond less rapidly after reinforcement than nonreinforcement—a tendency which is produced in the rat only by alternating reinforcement. The two patterns produced identical resistance to extinction in the fish, although the rat shows substantially greater resistance after random than after alternating reinforcement.

After a long series of extinctions separated by reconditioning sessions, the fish's resistance to extinction continued to vary with the amount of interpolated reinforcement. In a comparable experiment with the rat, resistance to extinction rapidly became independent of the amount of interpolated reinforcement.

VALUE OF BACKGROUND IN THE SPECIFICATION OF THE STIMULUS FOR JUDGMENT

By GLORIA ENGEL and ALLEN PARDUCCI, University of California, Los Angeles

The theory of *adaptation-level* was developed by Helson to provide a general treatment of judgment.[1] The theory states that each psychophysical judgment is determined by the physical ratio of the stimulus presented for judgment to the current level of adaptation. The adaptation-level, *AL*, is defined as the stimulus-value judged neutral (absolute judgment) or judged equal to the standard (comparative judgment), and the theory asserts that *AL* will be equal to a weighted mean of all stimuli affecting judgment. Thus each stimulus, whether singled out for judgment or merely present as background for the judged stimuli, pulls *AL* toward its own value just as each score in a distribution pulls the mean toward its own value.

The present studies investigate shifts in *AL* when background-stimuli are systematically varied. Application of Helson's theory to judgments of visual size suggests that *AL* should increase with increase in the size of the background or surround of the stimuli (*e.g.* the screen against which the stimuli are projected).[2] Thus, as the background becomes larger, the stimulus should be judged smaller since the ratio of its physical size to *AL* is smaller. This implication of *AL* theory is consistent with the results of several experiments.[3]

But how would an increase in the size of the background for some of the members of a series of stimuli affect the judgments of other stimuli

* Received for publication August 31, 1960. This work was aided by a grant, G7485, from the National Science Foundation.

[1] Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29; Adaptation level theory, in Sigmund Koch (ed.), *Psychology: A Study of Science*, 1, 1959, 565-621; Harry Helson, W. C. Michels, and Artie Sturgeon, The use of comparative rating scales for the evaluation of psychophysical data, this JOURNAL, 67, 1954, 321-326.

[2] The term *background* has elsewhere been used to characterize a number of different features of the stimulus-context, past or present. Throughout this paper, *background* refers specifically to additional features of the stimulus which are physically present at the same time as the stimulus singled out for judgment.

[3] Examples for judgments of visual magnitude: T. M. Künnapas, Influence of frame size on apparent length of a line, *J. exp. Psychol.*, 50, 1955, 168-170; Irvin Rock and Sheldon Ebenholtz, The relational determination of perceived size, *Psychol. Rev.*, 66, 1959, 387-401.

in the series for which the background was not varied? Insofar as this increase raised *AL throughout* the series, all of the judgments would be lowered. However, if the effect of each background is primarily upon the perceptual magnitude of the stimulus for which it is the immediate background, the judgments of the remaining stimuli might instead be raised. This follows from the possibility that the general *AL* would be pulled down by the lower (perceived or judged) values of the stimuli with the larger background. The *hypothesized direction* of the effect of a shift in background upon *AL* would thus depend upon whether the background was conceived as operating as a stimulus, *i.e.* in the same way as the judged stimuli, or instead as determining the contributions of the judged stimuli to *AL*. The first of the following experiments, using the *method of single stimuli,* investigates variation in *AL* as a function of variation in the background of only one of the stimuli in the series presented for judgment. The second experiment, using the *method of constant stimulus-differences,* studies the effects of independent variation of the backgrounds for the standard and comparison-stimuli.

Experiment I

Method. Thirty students from the course in introductory psychology were used as *Ss* under the three alternative experimental conditions, 10 *Ss* being exposed to each condition. All *Ss* were individually instructed to judge the size of each of a series of solid black squares they would see monocularly through a headrest-viewer. Judgments were to be announced aloud by *S* in terms of a five-category scale (*very small, small, medium, large,* and *very large*). The successive squares were projected from black-white transparencies onto a screen 3 ft. from *S* and surrounded by a baffle to reduce reflected light. Each square was exposed for approximately 3 sec., with 3 sec. between successive exposures. The presentations were divided into blocks of 10, consisting of 2 presentations of each square in randomized order. There were 5 such blocks or a total of 50 presentations under each condition.

The projected widths of the squares were 2.5, 4.0, 6.1, 9.1, and 11.2 cm., respectively, under all three conditions. The immediate background was white, with an outside width of 16.0 cm. for the four smaller squares under all three conditions and also for the largest square under the *Control* (C) condition; the background for this 11.2-cm. square was 13.9 cm. for the *Small* (S) and 19.5 cm. for the *Large* (L) conditions. The headrest-viewer restricted the visual field to the area of the largest background. Thus the only difference among the three conditions was in the size of the background for the largest square. Since it is known that the apparent size of the square would vary inversely with the size of its background, the purpose of the experiment was to determine how the judgments of the remaining squares would be affected by this variation.

Results. The data from the first block were discarded since the *Ss*

would not have seen all the squares before making their initial judgments, and their judgment-scales would not yet have stabilized. The largest square was judged *very large* on each of the remaining presentations by 29 of the 30 *S*s (one *S* in Condition C judging it *large* twice). There were, however, marked differences among the groups with respect to the judgments of the four smaller squares. These judgments were first converted to numerical form (*very small* = 1, *small* = 2, *medium* = 3, *large* = 4, and *very large* = 5); each *S*'s mean judgment of the four smaller squares was then computed, yielding a mean value based upon all the data for the stimuli common to all three conditions. For the three groups, the means and *SD*s, respectively, were as follows: 2.92 and 0.33 for L, 2.62 and 0.22 for C, 2.40 and 0.15 for S. An analysis of variance indicated that the variation in background had significant effects ($F = 12.67$, $df. = 2/27$, $p < 0.001$). The means and *SD*s of the *AL*s (each *S*'s *AL* being the geometric mean of the widths of the squares he judged *medium*) were as follows: 5.11 and 0.96 for L, 5.92 and 0.58 for C, 6.36 and 0.69 for S. This variation, based on a much smaller portion of the data, was also significant ($F = 6.66$, $df. = 2/27$, $p < 0.005$). The results thus indicate that both the judgment-scale as a whole and its midpoint, *AL*, vary *inversely* with the size of the background of one of the stimuli in the regular series, even though the judgments of that stimulus were not included in the analysis.

Experiment II

Method. The largest square, 11.2 cm. in width, was the only one presented in Experiment II, serving as both standard and comparison-stimulus for judgment by a modified *method of constant stimulus-differences*. Again, the three alternative backgrounds (13.9, 16.0, and 19.5 cm.) were employed, each being used equally often for both the standard and comparison-presentations. The task, for each pair of presentations, was to judge whether the second square, the comparison-stimulus, appeared to be larger or smaller than the first square, the standard. Each of the 9 possible combinations of standard and comparison-backgrounds was presented 18 times to each *S*, the order of presentation being randomized in blocks of 9.

The physical arrangements for Experiment II were identical with those of Experiment I. Each square was presented for 3 sec., with 2 sec. between standard and comparison and 10 sec. between the comparison and the next standard. Seven new *S*s were drawn from the introductory class. Two of these *S*s were instructed that they should try to ignore the backgrounds which would vary in size and thus interfere with the accuracy of their judgments.

Results. Each *S*'s percentage of *larger* judgments was computed for each of the nine conditions. Since the percentages for the two *S*s given the special instructions did not differ markedly from the rest, only the

mean percentages for all seven *S*s are presented (Table I). No statistical test would appear to be necessary. For the six combinations in which the standard and comparison-squares were different, all seven *S*s judged the square with the smaller background to be larger for a majority of the presentations. This distribution of judgments occurs in spite of the generally positive time-order error indicated both by the over-all mean, 46.8, and also by the mean percentages for the three conditions in which the standard and comparison-backgrounds were equal: 53.6, 40.3, and 37.7. While the experiment was not designed to determine the *AL* or *PSE* directly, the conclusion is clear that this value *increased* with increase in

TABLE I

MEAN PERCENTAGE OF JUDGMENTS OF 'LARGER' AS A FUNCTION OF WIDTH OF BACKGROUND

Background of standard square	Background of comparison square			
	13.9 cm.	16.0 cm.	19.5 cm.	Mean
13.9 cm.	53.6	4.4	2.3	20.1
16.0 cm.	97.4	40.3	17.0	51.6
19.5 cm.	97.1	71.0	37.7	68.6
Mean	82.7	38.6	19.0	46.8

the size of the background of the comparison-stimulus but *decreased* with increase in the standard's background size.

DISCUSSION

The results of these two experiments indicate that for both experimental procedures, comparative and absolute, *AL* may vary *inversely* with variation in the value of the backgrounds for the stimuli presented for judgment. This finding is in apparent contradiction to the general assertion of adaptation-level theory that each stimulus pulls *AL* toward its own value. Instead, the results suggest that each stimulus-value in the *AL* equation should be so rescaled as to allow for the contribution of its own background to its perceptual effects.

Helson's original equation for absolute judgment reads:

$$AL = \overline{V}^{a}B^{1-a},$$

where *AL* is the physical value (*e.g.* measured in cm.) of the stimulus judged medium, $\overline{V}$ is the geometric mean of all the stimulus-values presented for judgment, *B* is their background, and *a* is an empirical constant representing the relative influence of the judged stimuli and their background (such that $0 \leq a \leq 1$).[4] In the spirit of Helson's treatment of

[4] Helson, *op. cit.*, 1959, 584. In presenting Helson's own equations here, substitutions have been made for some of the symbols to simplify exposition in the

absolute judgment, it is here proposed that this equation be modified so that the background is treated not as a separate stimulus but as a partial determinant of the perceptual effect of the stimulus singled out for judgment. Specifically, it is proposed that the physical value of each stimulus (*e.g.* its width in cm.) be divided by a weighted geometric mean of its own value and the physical values of any other features of the stimulus-field (*e.g.* the width of the background in cm.) which affect its perceptual magnitude. Thus, the modified equation would read:

$$AL/AL^{a}B_{AL}{}^{1-a} = \overline{V}/\overline{V}^{a}\overline{B}_{V}{}^{1-a},$$

which simplifies, algebraically, to:

$$AL = B_{AL}\overline{V}/\overline{B}_{V},$$

where B_{AL} is the background for the stimulus judged medium and $\overline{B}_V$ is the geometric mean of the backgrounds for each of the stimuli in the series. In this form, the equation asserts that the physical magnitude of the stimulus judged medium, *AL*, may vary markedly with the physical magnitude of its own background, even after a long series of presentations (*i.e.* even after $\overline{V}$ and $\overline{B}_V$ have become relatively stable). However, the perceptually scaled value of *AL*, $AL/AL^{a}\, B_{AL}{}^{1-a}$, does stabilize in accordance with Helson's basic concept of an averaging process. Thus *AL* retains its operational meaning as the physical value of the attribute judged medium, but the concept of a general level of adaptation is better represented by the ratio,

$$AL/AL^{a}B_{AL}{}^{1-a}.$$

This modified equation is consistent with the finding in Experiment I that *AL* varies inversely with $\overline{B}_V$ (note that B_{AL} was not varied for the judgments from which the empirical *AL*s were computed since all *S*s in the L and S Conditions of Experiment I always judged the largest square *very large*). In addition, the new equation implies that an increase in the background of the stimuli judged medium would increase *AL* insofar as B_{AL} increased more than $\overline{B}_V$. It also implies that if all the backgrounds were the same, *AL* would be independent of their value.

With only three empirical *AL*s, the present data do not provide an adequate basis for a more quantitative evaluation of this proposed modification. The differences between the obtained *AL*s are larger, however, than would be predicted from the modified *AL* equation. This apparent discrepancy is consistent with the position, presented elsewhere, that *AL*

present context. A third term, representing the residual effects such as past experience, has been left out of the equation since this variable was not manipulated in the present experiments.

is a weighted mean of the *median* and the *midpoint,* or logarithmic mean of the two end-values of the presented stimuli.[5] Since the median stimulus is identical under all three conditions, substitution of a median-midpoint average for the values of $\overline{V}$ and $\overline{B}_V$ could provide a closer fit to the data. Simple substitution of the respective V and B midpoints provides a reasonably good fit which suggests that the end-values may require the major weighting under the present experimental conditions.

Helson's basic equation for *comparative* judgment is as follows:

$$AL = S^s \overline{V}^{1-s},$$

where $AL = PSE,$ the stimulus judged equal to the standard, S is the standard, $\overline{V}$ is the geometric mean of the comparison or variable stimuli, and s is the empirical weighting of the standard such that $0 \leq s \leq 1$.[6] Using the same rescaling of each stimulus-value in terms of background stimuli, this equation would read:

$$AL/AL^a B_{AL}{}^{1-a} = (S/S^a B_S{}^{1-a})^s (\overline{V}/\overline{V}^a \overline{B}_V{}^{1-a})^{1-s},$$

where B_{AL} is the value of the background for the stimulus judged equal to the standard, Bs is the background for the standard, and $\overline{B}_V$ is the geometric mean of the backgrounds for the variable stimuli. This simplifies algebraically to:

$$AL = B_{AL}(S/B_S)^s(\overline{V}/\overline{B}_V)^{1-s},$$

which correctly describes the following results of Experiment II: (a) that *PSE* varies *inversely* with the value of the standard's background, and (b) that *PSE* varies *directly* with the value of its own background. It also implies that when all the comparison-stimuli have the same background (*i.e.* $B_{AL} = \overline{B}_V$), *PSE* varies directly with the value of $B_V{}^s$; but if the standard also has the same background, *PSE* is independent of the value of the background.

The modified equation for comparative judgment may also be applied to situations in which the judgments are *completely* determined by the stimuli present at the time of judgment. In this simpler case, $s = 1$, and the ratio between the variable stimulus judged equal to the standard, *AL,* and its background must equal the ratio between the standard and the standard's background. This simplified relationship, $AL = B_{AL}\ S/B_S$, correctly describes the results of Wallach's experiment on brightness-

[5] Allen Parducci, R. C. Calfee, L. M. Marshall, and L. P. Davidson, Context effects in judgment: Adaptation level as a function of the mean, midpoint, and median of the stimuli, *J. exp. Psychol.,* 60, 1960, 65-77.

[6] Helson, Michels, and Sturgeon, *op. cit.,* 324.

constancy in which the intensities of two, simultaneously-presented pairs of ring-disks were independently varied.[7]

While the modifications proposed here somewhat complicate Helson's basic *AL*-equations, they do not seem to represent a radical departure from his general position. His continued emphasis upon the geometric mean, a provisional commitment to Fechner's logarithmic scaling of the stimuli, suggests that the stimuli are to be scaled in accordance with their perceptual effects. But since we know that the immediate perceptual response to each stimulus is largely determined by the physical relationship between stimulus and background, it seems that Helson's position should require that this relationship be incorporated into the definition or scaling procedure for each stimulus value.[8] By the same line of reasoning, the physical values of other stimuli (*e.g.* those presented on immediately preceding trials) might also be incorporated into the specification of each stimulus, but that would perhaps involve an impracticable complication of the equations for adaptation-level.

Summary

Both absolute and comparative judgments of size were shown to vary systematically with variation in the size of the background for projected squares. The direction of some of the shifts in judgment was in apparent contradition with previous statements of adaptation-level theory. It was proposed that the theory be modified so that the value of each stimulus in the equations for adaptation-level would be defined with respect to a weighted mean of the stimulus and its simultaneously presented background.

[7] Hans Wallach, Brightness constancy and the nature of achromatic colors, *J. exp. Psychol.*, 38, 1948, 310-324.

[8] A similar suggestion has previously been made with respect to theories of discrimination-learning by D. A. Riley, The nature of the effective stimulus in animal discrimination learning: Transposition reconsidered, *Psychol. Rev.*, 65, 1958, 1-7.

CONTEXT-EFFECTS IN JUDGMENTS OF LENGTH

By ALLEN PARDUCCI and LOUISE M. MARSHALL,
University of California, Los Angeles

Research on judgment has often involved experimental manipulation of the stimulus-context in which simple perceptual judgments are made. It is a commonplace that all judgments are relative; the problem has been to discover what they are relative to—that is, to specify how different features of the context determine the judgments. This has been the central concern of the theory of *adaptation-level* which asserts that the category-scale for psychophysical judgments is centered at the weighted mean of all stimuli affecting the judgments.[1] Recent research has indicated, however, that the mean may not be the most useful parameter for characterizing the stimulus-context. Independent manipulation of either the *midpoint* (the mean of the two end-stimuli) or the *median* of the set of stimuli presented for judgment had significant effects upon the centering of the scale of judgment, but manipulation of the mean appeared to have no effect when midpoint and median were held constant.[2] These findings were interpreted as consistent with the hypothesis that the scale of judgment reflects a *compromise* between two different tendencies: (1) to divide the range of stimuli presented for judgment into proportionate sub-ranges; and (2) to use the alternative categories of judgment with proportionate frequencies.

This interpretation was based upon experiments studying absolute judgments of numerical magnitude. The present research was designed to investigate the same variables with a more clearly perceptual dimension, length of line. In other respects, the basic logic of the research remained unchanged. Thus, *S*s made absolute judgments of the lengths of different sets of lines. The sets differed with respect to the physical values of their midpoints, medians, and means. The primary purpose was to determine how the scale of judgment is affected by variation of these parameters of

* Received for publication August 31, 1960. This paper reports part of a series of investigations supported by a grant, G7485, from the National Science Foundation.

[1] Harry Helson, Adaptation level theory, in Sigmund Koch (ed.), *Psychology: A Study of Science,* 1, 1959, 565-621; D. M. Johnson, *The Psychology of Thought and Judgment,* 1955, 340-348.

[2] Allen Parducci, R. C. Calfee, L. M. Marshall, and L. P. Davidson, Context effects in judgment: adaptation level as a function of the mean, midpoint, and median of the stimuli, *J. exp. Psychol.,* 60, 1960, 65-77.

the sets of stimuli. A second objective was to determine how the relative effects of these parameters vary with the mode of presentation of the stimuli. The interpretation of judgment as a range-frequency compromise led to the expectation that the median would have greater influence when presentation stressed the relative frequencies of the different lengths of line than when the differences in length between lines were emphasized.

METHOD

Stimuli. Eleven different sets of 45 lines each were used. The lines in each set were printed parallel to the short sides of an 8 × 11-in. page. Under all conditions, the successive lines, each approximately 0.5 mm. in width, were ordered with respect to length; the longest line was printed 30 mm. from the top, the shortest 20 mm. from the bottom, and each line ended 15 mm. from the right side of the page. Table I shows the lengths of the lines in the 11 sets. The differences between the lengths of successive lines for a given set were approximately equal within each of the 20-mm. intervals indicated in Table I. Table II shows how the midpoints of the lengths differ for the sets with the low midpoint (*LMP*) and high midpoint (*HMP*), the medians for the sets with the low median (*LMD*) and high median (*HMD*), and the means for the sets with the low mean (*LM*) and high mean (*HM*). It should be noted that for each of these three pairs, the other two parameters are constant at 100 mm., which is the value of the midpoint, the media, and the mean for the rectangular (*RECT*) set. The four remaining sets of lines, *LL, LH, HL,* and *HH,* form a 2 × 2 factorial design, with midpoint and median as the independent variables.

TABLE I

NUMBERS OF LINES IN DIFFERENT INTERVALS OF THE RANGE OF LENGTH

Stimulus-set	Interval of length (mm.)								
	10–29	30–49	50–69	70–89	90–109	110–129	130–149	150–169	170–190
RECT	5	5	5	5	5	5	5	5	5
LMP	2	2	2	12	9	6	12	0	0
HMP	0	0	12	6	9	12	2	2	2
LMD	2	2	2	17	10	2	2	2	6
HMD	6	2	2	2	10	17	2	2	2
LM	11	3	3	3	5	11	3	3	3
HM	3	3	3	11	5	3	3	3	11
LL	2	2	5	13	5	5	7	6	0
LH	5	3	3	3	8	12	7	4	0
HL	0	4	7	12	8	3	3	3	5
HH	0	6	7	5	5	13	5	2	2

Each of the sets was presented with two different spacings between the lines. For *frequency-spacing,* all lines were separated by the same vertical distance on the page, 5 mm. For *length-spacing,* the vertical separation between the lines varied in direct proportion to the difference between their lengths so that each set of lines formed a trapezoidal pattern with the left ends falling on a direct line between

the left ends of the longest and shortest lines. It should be noted that the requirements of both kinds of spacing were thus met by the same arrangement of lines for the *RECT* set.

Procedure. The experiment was performed in sections of introductory psychology. Copies of each of the sets of lines were dittoed and stapled to a cover page, on which the following instructions appeared:

> On the next page, there is a set of lines placed in order of decreasing length. Your job is to study the entire set and then to decide how long or short each line is—in comparison with all the other lines on the page. Record to the right of each line your judgment of its relative length. For each line, use one of the following categories of length: *very long*, 6; *long*, 5; *slightly longer than average*, 4; *slightly shorter than average*, 3; *short*, 2; *very short*, 1. Thus if a line seems *short* (compared to the other lines), write a small '2' beside its extreme right end; if you judge it to be *very long* (by comparison), record a '6.' Be sure to write one of the six numerals to the right of each line. Take your time with these judgments, but do not try to devise any special rules for judgment. The purpose of this research is to find out how people, in general, compare the lengths of different lines.

After oral instructions to work independently, the materials were distributed in a randomized order, one set of lines for each student. All materials were collected after approximately 10 min. There was some inequality in the numbers of *Ss* receiving the various sets, and the data from some *Ss* were discarded because of failure either to judge all of the lines or to maintain the numerical order of the categories. For subsequent analysis, all groups were reduced by random elimination to 27 *Ss*, except for the *RECT*, *LMP*, and *HMP* frequency-spacing conditions ($N = 34$ each).

RESULTS

Analyses of variance. The primary dependent variable was the adaptation-level (*AL*), defined as the midpoint between the longest line judged '3' and the shortest line judged '4' (*i.e.* the arithmetic mean of these two lengths). The means and *SD*s of the *AL*s for each of the 21 experimental conditions are presented in Table II. Separate analyses of variance were performed upon the *AL*s for three sub-sets of these conditions to determine the significance of the effects of independent manipulation of the midpoint, median, and mean, respectively, and the interaction of each of these variables with the spacing of the lines upon the page (*i.e.* a series of 2×2 analyses).

The effect upon *AL* of variation in midpoint, *LMP vs. HMP*, was highly significant ($F = 15.23$, $df. = 1/118$, $p < 0.001$), as was the effect of variation in median, *LMD vs. HMD*, ($F = 57.00$, $df. = 1/104$, $p < 0.001$). The interaction between median and spacing was also significant in this analysis ($F = 16.48$, $df. = 1/104$, $p < 0.001$), with the effects of variation in median being greater with frequency-spacing. Variation in mean, as determined for the *LM* and *HM* Sets, did not have significant

effects; however, the effects of this variation interacted with spacing ($F = 4.80$, df. $= 1/104$, $p < 0.05$).

The $2 \times 2 \times 2$ factorial analysis of the *AL*s for the *LL*, *LH*, *HL*, and *HH* Sets, with each kind of spacing, indicated significant effects of the median ($F = 51.50$, $df. = 1/208$, $p < 0.001$) and also a significant interaction between midpoint and stimulus spacing ($F = 5.66$, $df. = 1/208$, $p < 0.025$). None of the other sources of variance, either in this or in any of the three preceding analyses, was significant at the 5% level. In summary, significant differences in *AL* were found associated with variation in either the midpoint or the median but not with the mean. The effects of

TABLE II

STIMULUS-PARAMETERS AND MEAN ADAPTATION-LEVELS

Stimulus-set	Stimulus-parameters			Frequency-spacing		Length-spacing	
	median	midpoint	mean	*AL*	*SD*	*AL*	*SD*
*RECT**	100	100	100	106.0	10.0	106.0	10.0
LMP	100	80	100	97.8	12.7	95.6	11.5
HMP	100	120	100	103.9	6.6	104.4	9.5
LMD	88	100	100	90.9	5.5	97.5	11.5
HMD	112	100	100	110.8	6.9	103.4	10.3
LM	100	100	89	104.5	9.3	106.1	12.3
HM	100	100	111	106.4	10.3	99.4	8.8
LL	92	90	100	97.3	8.1	93.8	9.7
LH	110	90	100	108.8	10.5	103.6	10.1
HL	90	110	100	94.6	9.2	100.5	11.4
HH	108	110	100	107.1	6.9	105.2	11.8

* The two spacing-conditions are identical for *RECT*, so that *AL* and *SD* are based upon the responses of a single *RECT* group.

each of these three parameters interacted significantly with the spacing of the stimuli.

Multiple-regression analyses. Mean *AL* was used as the criterial variable in two separate multiple-regression analyses, one for each kind of spacing, with the midpoints and medians of the sets as the predictor variables. The least-squares equation is AL = 0.86(Median) + 0.14(Midpoint) for frequency-spacing; for length-spacing, $AL = 0.64(\text{Median}) + 0.36(\text{Midpoint})$. The relative weighting of the median is greater for frequency-spacing, and this is consistent with the analyses of variance. As compared with the results for judgments of numerical magnitude, where $AL = 0.45(\text{Median}) + 0.55(\text{Midpoint})$,[3] the midpoint had relatively less influence upon the judgments of length, regardless of spacing. The unbiased multiple correlations for frequency- and for length-spacing were

[3] Parducci, Calfee, Marshall, and Davidson, *op. cit.*, 74.

0.85 and 0.73, respectively. Both the median-midpoint combinations accounted for a significant portion of the variance in *AL* ($F = 31.6$, $df. = 2/9$, $p < 0.001$ for frequency-spacing; for length-spacing, $F = 7.5$, $p < 0.025$).

Frequencies and widths of categories. The previous analyses were based upon the value of the middle limen, *AL*. While this measure tends to correlate positively with each of the other four limens, additional information about the effects of stimulus-spacing can be obtained through a comparison of the numbers of lines and the ranges of lengths included in each of the categories of judgment. A tabulation was made of the mean number of lines placed in each of the six categories under each of the 21 experimental conditions. These values for the *RECT* condition were (in order from *very long* to *very short*): 5.7, 7.4, 8.0, 9.2, 8.4, and 6.3. The slight asymmetry (more lines in the 'shorter' categories) is consistent with the ordinal effects previously reported for similar situations if one assumes that *S*s tend to start at the top of the page and work down toward the shorter lines at the bottom.[4]

Two difference-scores were computed for each condition in order to measure the changes in the scale of judgment associated with the differences in stimulus-spacing. The first of these difference-scores was simply the mean of the six differences (taken without regard to algebraic sign) between the mean number of lines in each of the six categories under a given condition and the corresponding *RECT* values. The resulting score indicates how similar the scale of judgment for each condition was to the *RECT* scale with respect to frequency of use of the different categories. A second measure was obtained for each condition by first determining the mean percentage of the range covered by each of the six categories and then computing the mean absolute difference between these six values and the corresponding mean percentages for the *RECT* condition. This second difference-score indicates how similar the judgment-scale for each experimental group was to the *RECT* scale with respect to the relative width or range of the respective categories.

The mean absolute difference in frequency per category, averaged over all 10 stimulus-sets, was 1.3 lines for frequency-spacing and 2.8 lines for length-spacing, or 2.9% and 6.2%, respectively, of the 45 lines. For each of the 10 stimulus-sets, the mean deviation from the *RECT* baseline was at least ⅔ greater with length-spacing. The mean absolute difference in the widths of the categories was 4.4% for frequency-spacing but only

[4] Parducci, Ordinal effects in judgment, *J. exp. Psychol.*, 58, 1959, 239-246; Parducci, Calfee, Marshall, and Davidson, *op. cit.*, 70.

2.7% for length-spacing, the direction again holding for all 10 sets. These comparisons show that the tendency to use the various categories of judgment with proportionate frequencies (the proportions indicated by the frequencies in *RECT*) is stronger with frequency-spacing and that the tendency to divide the range into proportionate sub-ranges (again following the proportions in *RECT*) is stronger with length-spacing. Thus, the differences in strength between these hypothetical judgmental tendencies, apparent for *AL*, also appear when the measure is based upon all six categories of judgment.

Perceptual grouping. The arrangement of the lines, particularly in some of the sets with length-spacing, produces perceptual groupings. In such cases, category-limens tend to fall at the boundary of a group of equally spaced lines. Thus for the *LM* set, almost half the *S*s given length-spacing broke their *very short* category just above the 29-mm. line. As can be inferred from Table I, the lines were closely spaced between 10 and 29 but relatively sparse between 30 and 109 mm. The 11 closely-spaced lines between 110 and 129 mm. also appeared as a separate group, constituting a single judgment-category for almost half the *S*s judging the *LM* set with length-spacing. This kind of grouping was less pronounced for the frequency-spacing conditions, but even with the equal vertical spacing there appeared to be a slight tendency for the category-limens to fall at points where the slope formed by the left ends of the lines changed dramatically.

It would be difficult to disentangle the effects of this perceptual grouping from the various independent variables with respect to which the judgments have been analyzed. Any evaluation of the grouping is *post hoc* since the stimulus-sets were not selected to permit independent manipulation of the grouping factors. For conditions in which grouping was pronounced, separate tabulation was made, however, of the *AL*s of *S*s who did not break their categories at the borders of the groups. The differences between the mean *AL*s for these *S*s and the mean *AL*s for the remaining *S*s were small enough to suggest that grouping tendencies did not materially affect any of the major conclusions of the experiment.

DISCUSSION

These perceptual groupings share with other phenomena of judgment a general dependence upon the physical relationships between the stimuli presented for judgment. The limens or boundaries of these groupings occurred at points of abrupt shift, either in the average vertical distance between successive lines (length-spacing) or in the slope formed by the

left ends of the lines (frequency-spacing). Insofar as *S*s divide the range into proportionate sub-ranges, limens would be more likely to occur at the points of greatest difference in length between successive lines. It should be noted that the grouping-limens occurred instead at the points of greatest shift in the magnitude of successive differences, averaged over a sequence of stimuli. The specification of the stimulus relationships which determine this kind of perceptual grouping might itself become an objective for research. The concern of the present analysis was, however, that this unforeseen grouping did not obscure the effects upon judgment of those experimental conditions which were systematically manipulated.

The effects of the major experimental conditions upon judgments of length are clearly consistent with the results previously obtained for judgments of numerical magnitude. For both length- and frequency-spacing, *AL* varied directly with independent variation of either the midpoint or the median. There appeared to be no systematic relationship between *AL* and the mean across conditions of spacing.

Relationship to range-frequency compromise. The present results carry the argument beyond the conclusions of the research on numerical magnitude by showing that the relative weighting of the midpoint and median can be systematically manipulated in accordance with the interpretation of judgment as a range-frequency compromise. The alternative methods of spacing the stimuli were chosen to emphasize either the differences in the number of lines placed in the various categories by *S*s or the differences in the widths of their categories. In both cases, the emphasis consisted of correlating these differences with differences in vertical position upon the page. When the lines were equally spaced so that *S*s could more readily determine the relative number of lines they were placing in each category, their category-frequencies conformed more closely to those of the *RECT* group and their *AL*s correlated more closely with the medians of the stimulus sets. When the vertical distance between the lines was in direct proportion to the difference in their length, the proportions of the total range included in the various categories conformed more closely to the corresponding *RECT* proportions and the *AL*s correlated more closely with the midpoints of the stimulus sets.

Relationship to the theory of adaptation-level. The relationship between *AL* and the midpoint might be incorporated into Helson's weighted-mean analysis by weighting the two end-stimuli more heavily than the other stimuli in the set presented for judgment. It would be much more difficult to specify how the other stimuli are to be weighted in order to account for the effects of variation in the median. And insofar as *AL* is

unaffected by any variation of stimuli which leaves the midpoint and median unchanged, the stimuli which were varied would have zero weighting (*i.e.* they would have no place in the adaptation-level formula). This appears inconsistent with the notion of stimulus-interaction or pooling which is basic to the theory of adaptation-level.[5]

As emphasized elsewhere, the mean of a set of stimuli is often intermediate in value between the midpoint and median.[6] When this is the case, the mean may provide a satisfactory first approximation for the determination of *AL*. The *AL*s for two sets of stimuli may, however, be identical and yet their other category-limens quite different. Thus the *AL*s for the *RECT* and *HM* sets are very close with frequency-spacing, but the *HM* limen between *very short* and *short* was 12 mm. longer (a highly significant difference). Since there were fewer lines within the lower third of the *HM* range (see Table I), this limen would be higher for *HM* than for *RECT* insofar as *S*s tended to judge the same number of lines *very short.* Corresponding differences, in each case consistent with the compromise-interpretation, occur at the other category-limens for these two conditions. Thus, even when *AL*s are equal, the range-frequency compromise may produce marked differences in other parts of the scale of judgment.

SUMMARY

Sets of lines were presented for judgment in terms of absolute categories of length. The sets differed with respect to the midpoint, median, and mean lengths of the lines and also with respect to the spacing between the lines. The center of the scale of judgment, *AL*, was found to vary systematically with variation in either the midpoint or median but not with the mean. The relative effect of variation of the median was greater when the stimuli were spaced so as to stress differences in frequencies rather than differences in length. The results were interpreted as consistent with the representation of judgment as a compromise between the tendency to divide the stimulus-range into proportionate widths and to use the judgment-categories with proportionate frequencies.

---

[5] Helson, *op. cit.*, 569; Johnson, *op. cit.*, 344.
[6] Parducci, Calfee, Marshall, and Davidson, *op. cit.*, 77.

SIMULTANEOUS AND SUCCESSIVE DISCRIMINATION-REVERSAL IN THE RAT

By R. C. GONZALEZ, Bryn Mawr College, and
BRYAN SHEPP, University of Connecticut

A discriminative problem of the so-called *simultaneous* type provides differential reinforcement in at least one afferent dimension; its solution can be described in terms of the development of excitatory properties by the positive stimulus and of inhibitory properties by the negative stimulus.[1] A problem of the so-called *successive* type cannot be analyzed in the same way, because it provides no differential reinforcement in any afferent dimension. Consider, for example, a successive black-white discrimination with either two black or two white stimulus-objects on each trial, in which, say, the right-hand black and the left-hand white are designated as positive. In such a problem, black is as often reinforced as white, and leftness is as often reinforced as rightness. With no one of the four components reinforced more than any other, how is solution of the successive problem possible? Two accounts have been proposed: *(1)* one based on the notion of *conditional discrimination* or *stimulus-compounding*;[2] and *(2)* the other based on the concept of *configurational discrimination*.[3]

(1) According to the compounding interpretation, the successive solution is based on the acquisition of excitatory and inhibitory properties by *combinations* of stimulus-components. Certain visual-spatial compounds (*e.g. black-right* and *white-left*) are assumed to acquire excitatory properties, while others (*e.g. black-left* and *white-right*) acquire inhibitory properties; that is, the compounds are assumed to function in a manner which cannot be derived by summation from the functions of the components. Spence has asserted that the successive solution is more difficult than the simultaneous, and that compounding will develop only when a component-solution is not possible.[4]

---

* Received for publication April 10, 1960. This experiment was conducted in the Laboratory of Psychopharmacology at the University of Maryland.

[1] K. W. Spence, The nature of discrimination learning in animals, *Psychol. Rev.*, 43, 1936, 427-449.

[2] Spence, The nature of response in discrimination learning, *Psychol. Rev.*, 59, 1952, 89-93; H. W. Nissen, Description of the learned response in discrimination behavior, *Psychol. Rev.*, 57, 1950, 121-131.

[3] Philip Weise and M. E. Bitterman, Response-selection in discriminative learning, *Psychol. Rev.*, 58, 1951, 185-195; Bitterman, Spence on the problem of patterning, *Psychol. Rev.*, 60, 1953, 123-126.

[4] Spence, *op. cit.*, 89-93.

(2) According to the configurational interpretation, the successive solution may be achieved in terms of certain unitary properties of the *pairs* of stimuli, the animal learning simply to make opposed responses to the two pairs (*e.g.* go right to the configuration *black-black* and left to the configuration *white-white*). From this point of view, the animal discriminates, not between simultaneously presented compounds, but between successively presented configurations. Furthermore, no functional priority is assigned to the component-solution. A successive problem may be mastered before a corresponding simultaneous problem, and a configurational solution may be achieved even where a component-solution should be possible.[5]

Experimental tests of these divergent conceptions of the successive solution have centered primarily on the initial difficulty of simultaneous and successive discrimination. Initial difficulty, however, is markedly influenced by situational variables, the most important of which are *contact* and *contiguity*.[6] Under conditions of indirect approach to spatially contiguous stimuli (*noncontact, contiguity*), the simultaneous problem is more difficult than the successive. Under conditions of direct approach to spatially separated stimuli (*contact, noncontiguity*), and under more conventional conditions, such as the *T*-maze or the two-window jumping apparatus (*contact, contiguity*), the successive problem is more difficult than the simultaneous. Marked shifts in relative difficulty occur with practice. When *S*s learn two consecutive problems of each type, transfer to the second problem is much greater for *S*s having learned the initially more difficult one, and in the second problem the difference in difficulty between simultaneous and successive presentations tends to disappear.[7] This finding suggests that the effects of situational variables on relative difficulty are transient, and that an extended series of discriminations of each type would weight intrinsic differences in difficulty at the expense of contextual ones.

In the present experiment, the work on transfer was extended to the study of a long series of consecutive discriminations in the context of a habit-reversal design. Two groups of rats, one presented with a simultaneous and the other with a successive problem, each were given one day of

[5] Weise and Bitterman, *op. cit.*, 185-195; D. C. Teas and M. E. Bitterman, Perceptual organization in the rat, *Psychol. Rev.*, 59, 1952, 130-140; J. R. Turbeville, A. D. Calvin, and M. E. Bitterman, Configurational and relational learning in the rat, this JOURNAL, 65, 1952, 424-433.

[6] For a review of the factors affecting relative difficulty of the two types of problem, see M. E. Bitterman, D. W. Tyler, and C. B. Elam, Simultaneous and successive discrimination under identical stimulating conditions, this JOURNAL, 68, 1955, 237-248.

[7] Jerome Wodinsky, M. A. Varley, and M. E. Bitterman, Situational determinants of the relative difficulty of simultaneous and successive discrimination, *J. comp. physiol. Psychol.*, 47, 1954, 337-340; M. E. Bitterman and J. V. McConnell, The role of set in successive discrimination, this JOURNAL, 67, 1954, 129-132.

training with vertically and horizontally striped cards in a two-window jumping apparatus, and the preference so established then was reversed daily for 32 days.

Method

Subjects. The Ss were 20 male albino rats of the CF Nelson Strain, 90-120 days old at the beginning of the experiment.

Apparatus. The apparatus employed was a conventional Lashley two-window jumping apparatus modified to eliminate punishment for errors. Narrow runways branched out from a starting platform 9 in. removed from the two windows, which were 5.5 in. square and situated 2.5 in. apart. The response of *S* was to run from the starting platform to one of the windows and to push against the stimulus-card; the incorrect card was locked in place, but the correct card gave access to a feeding platform in the rear.

Four black and white striped stimulus-cards were employed. Two of the cards were horizontally striped and the other two were vertically striped. The width of each stripe on each card was 0.5 in.

Preliminary training. Throughout the experiment the animals were maintained on a 24-hr. feeding schedule. After an initial period of handling, during which they were fed wet mash on the feeding platform, the animals were taught to run from the platform, first to open windows, and then to unlocked mid-gray cards. Manual guidance was used from the beginning to provide equal experience with both windows. At the conclusion of the preliminary training, the animals were divided into two groups of 10 *Ss* each, matched for position-preference and for adjustment to the experimental situation.

Experimental training. The animals of the simultaneous group were confronted with one vertically striped and one horizontally striped card on each trial of the experiment proper, the lateral position of the two cards being varied systematically in accordance with selected Gellermann-orders.[8] On the first day (Reversal O), 5 *Ss* were trained with the vertically striped card positive, and the remaining 5 with the horizontally striped card positive. The animals of the successive group were confronted with two vertically striped cards on half of each day's trials, and with two horizontally striped cards on the remaining half, the order of presentation following the same Gellermann-orders as those employed with the simultaneous group. On the first day of training, half the *Ss* of the successive group were rewarded for responding to the right-hand card of the vertically striped pair and to the left-hand card of the horizontally striped pair, while the direction of correct response was the opposite for the remaining animals. Both groups were reversed *daily* for a total of 32 reversals. That is, a simultaneous animal for which vertical stripes were positive in Reversal O was reinforced for response to horizontal stripes in Reversal 1, for response to vertical stripes in Reversal 2, and so forth. Similarly, a successive animal which was trained right to vertical stripes and left to horizontal stripes in Reversal O was rewarded for going left to vertical and right to horizontal in Reversal 1, and so forth. Throughout the experiment, there were 20 massed

---

[8] L. W. Gellermann, Chance order for alternating stimuli in visual discrimination experiments, *J. genet. Psychol.*, 42, 1933, 356-360.

trials per day by the correction method. Following an incorrect response S was returned to the starting platform and manually guided in the correct direction; a correct response admitted S to the feeding platform and terminated the trial.

RESULTS

The performance of the two groups, measured in terms of mean number of errors per reversal, is shown in Fig. 1. Each curve begins at the chance-level and shows a progressive decline, which seems greater in the

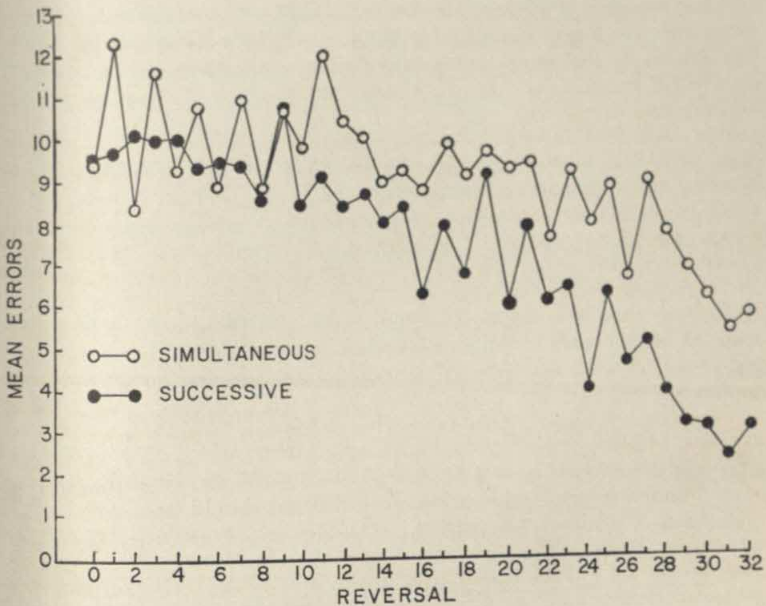


FIG. 1. PERFORMANCE OF THE TWO GROUPS DURING THE 33 DAYS OF THE EXPERIMENT

successive than in the simultaneous group, and this impression is substantiated by analysis of variance ($F = 35.99$ with 1 and 18 *df.*, $p < 0.01$). The variance due to reversals ($F = 15.92$ with 32 and 576 *df.*) and that due to the interaction of groups and reversals ($F = 2.16$ with 32 and 576 *df.*) also are significant beyond the 1% level of confidence.

Although the curves overlap in the early stages of training, marked differences in pattern are evident. Transfer-effects from one reversal to the next are prominent in the simultaneous group but negligible in the successive group. The difference between the two curves in the early reversals,

which is responsible in large part for the significant interaction, suggests greater learning in the simultaneous group. This result parallels the findings of previous experiments which show that, under conventional conditions, the simultaneous problem is easier than the successive. How, then, is the superior performance attained by the successive group in the final stages to be explained? The interpretation of Bitterman and Wodinsky, which served as the basis for the present experiment, provides an answer.[9]

These investigators proposed that, in any given situation, the experimental conditions employed tend to favor either the perception of the unitary properties of each pair of stimuli as a whole (configurational organization) or the perception of the unique properties of the individual stimuli (component-organization). On the assumption that successive problems are learned configurationally (on the basis of the difference *between-pairs*) and simultaneous problems in terms of components (on the basis of the differences *within-pairs*), the initial difference in speed of learning should be a function of the perceptual set fostered by the experimental conditions. Under configurational conditions (*noncontact, contiguity*), a set to respond configurationally is promoted from the outset, thus facilitating the successive solution and retarding the simultaneous solution. Component conditions (*contact, noncontiguity*) hinder the development of a configurational set, thus retarding successive solution and facilitating simultaneous solution. When, however, the transient effects of set are overcome (either by learning to some criterion-level or through extensive training such as provided in the present experiment), the essential simplicity of configurational organization should be manifested in superior successive performance on subsequent problems, irrespective of the experimental conditions.

The present experiment provides a systematic test of the validity of this principle. Under the conventional conditions employed, which favor the perception of components, initial learning was more rapid in the simultaneous group. With continued training this advantage disappeared and, finally, in the latter stages of training, the successive group gained a clear superiority. Taken together with presently available data, these results point to the necessity of distinguishing between two qualitatively distinct processes which operate in simultaneous and successive discrimination. The original emphasis of Weise and Bitterman on the functional priority of configurational as compared with component and relational organiza-

[9] M. E. Bitterman and Jerome Wodinsky, Simultaneous and successive discrimination, *Psychol. Rev.*, 60, 1953, 371-376.

tion seems justified,[10] but initial dominance of process is seen to depend upon the perceptual characteristics emphasized by the experimental situation.

SUMMARY

Two groups of rats, one trained on a simultaneous problem and the other on a successive problem were given a series of 32 daily reversals under conventional conditions in a modified jumping-apparatus. The typical initial superiority of simultaneous solution under these conditions was seen to be a function, not of an intrinsic difference in the difficulty of the two types of problem, but of the transient effects of set. Although the simultaneous group showed superior learning in the early stages of the experiment, the rate of improvement in reversing was greater in the successive group, as was its final level of performance. The results are interpreted as consistent with a configurational and as contradictory to a compounding interpretation of successive solution.

---

[10] Weise and Bitterman, *op. cit.*, 185-195.

A RE-EVALUATION OF EVIDENCE OF ONE-TRIAL ASSOCIATIVE LEARNING

By G. R. LOCKHEAD, IBM Research Center

Results of recent studies suggest that *Ss* form rote-associations in one trial. The authors of these studies argue that repetition of the material may help only to allow new associations to be formed on each trial and to strengthen those associations once they have been made.[1] This argument is contrary to the notion that associations are formed by gradual strengthening.

In the studies which support the hypothesis of one-trial learning, *S* was presented a list of paired associates, one pair at a time, which he was to learn. He was then shown the left-hand member of each pair and required to respond with the right-hand member. In the control group, this procedure was repeated until *S* could respond correctly to the entire list. In the experimental group, whenever *S* did not respond correctly to a left-hand member, that member and its response were replaced by a new pair. This revised list was then shown *S* in the next trial. A trial is the presentation of the list of pairs of syllables followed by the recall-test for all syllables. After the test-trial on the revised list, those pairs to which *S* did not respond correctly were replaced. If *S* failed on a syllable he had previously given correctly, that pair was also dropped. The procedure continued until *S* could respond correctly to an entire list. Thus, in the non-replacement condition *S* saw the same material over and over until he had learned it to criterion, while in the replacement-condition *S* had to respond correctly the first time, or he never saw that pair again. The material was reordered for each trial in both conditions and *S* was told 'right' or 'wrong' after each response, as noted in Rock and Heimer's article.[2]

In none of the one-trial studies was the non-replacement group superior to the replacement group. From this result, it has been concluded that as-

---

* Received for publication July 14, 1960. The author is indebted to E. T. Klemmer and R. Seibel for their invaluable suggestions.

[1] L. L. Clark, T. G. Lansford, and K. M. Dallenbach, Repetition and associative learning, this JOURNAL, 73, 1960, 22-40; Irvin Rock, The role of repetition in associate learning, this JOURNAL, 70, 1957, 186-193; Irvin Rock and Walter Heimer, Further evidence of one-trial associative learning, this JOURNAL, 72, 1959, 1-16; Michael Wogan and R. H. Waters, The role of repetition on learning, this JOURNAL, 72, 1959, 612-613.

[2] Rock and Heimer, *op. cit.*, 1959, 3; not, however, in the other studies.

sociations are formed not by a process of gradual strengthening, but in one trial.

In Rock's studies using this procedure, *S* was shown a pair to be learned for 3 sec. and then allowed a 5-sec. interval before the next pair appeared. This rate of presentation allowed a total of 8 sec. in which that pair could be rehearsed. Further, *S* was allowed 30 sec. between trials, while *E* reorganized the material, during which *S* could rehearse. Thus, although *S* had been given only one presentation of a particular pair, he had opportunity to rehearse it covertly a large number of times. Learning may, therefore, have been achieved by gradual strengthening, but this strengthening took place in an unobservable manner.

The present experiment is a replication of Rock's Experiment II with (1) a pair of syllables presented once every 8 sec., as in Rock's study; and (2) with the over-all rate of presentation increased by nearly a factor of four.[3] The faster rates chosen were as fast as would allow *E* to manipulate the materials accurately.

Method

Two groups of 14 and two groups of 10 college students were required to learn a list of eight paired nonsense syllables of 47% and 53% associative value according to Glaze.[4] These syllables were typed on plain white IBM cards using the extended spacing on an IBM Executive typewriter. *S* viewed the cards through a tachistoscope with an opening of 3 × 3 in. on a vertical plywood-board 30 × 20 in. in size, which prevented *S* from seeing how the material was being reordered. *S* was seated about 24 in. in front of the material and assigned alternately to replacement and non-replacement conditions, with only one rate of presentation being made on any one day.

For the fast rate, the tachistoscope allowed *S* to see the syllables for 0.75 sec. and a white uniform field, of the same intensity, for 1.25 sec. In the experiment which attempted to replicate Rock's study, these times were 3 and 5 sec. for the presentation of the pairs, and 2 and 3 sec. for the recall-test. This procedure differs from Rock's, which allowed 5 sec. for the recall-test, with no reported inter-syllable interval. The total time allowed for each recall-test was the same, however. Thus, the fast rate of presentation was higher by a factor of four for the presentation of the pairs and 2.5 for the recall-test over those rates used by Rock.

With the fast rate, there was a 6-sec. pause before the recall-test and a 12-sec. interval between trials. A 30-sec. rest-period followed every 10 trials during which *Ss* were encouraged to talk to limit rehearsal. The rest-period was intended to reduce some of the effects of massing. In the replication of Rock's study, the only time-delay was a 30-sec. pause between trials.

[3] Rock, *op. cit.*, 1957, 189.

[4] J. A. Glaze, The association value of nonsense syllables, *J. genet. Psychol.*, 35, 1928, 255-267. Cf. also E. H. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 541-544.

The *Ss* in the non-replacement conditions for both rates of presentation were shown a list of pairs of syllables, one pair at a time, and then a list of the left-hand members of each pair, again one at a time, to which they were to respond. This procedure was repeated until *S* had responded correctly to an entire list. The *Ss* in the replacement-conditions were presented the list of syllables by the same procedure as that used in the non-replacement condition with the exception that whenever *S* did not respond correctly to any syllable, that pair of syllables was replaced with a new pair on the next trial. There were 10 *Ss* in both conditions of the replication of Rock's study and 14 *Ss* in both conditions of the fast-rate study. In both conditions, *Ss* were told 'right' or 'wrong' immediately after each response.

To record responses and reorder the cards at the faster rate, two *Es* were necessary in the replacement-condition. Only one *E* was used for the non-replacement condition.

RESULTS

Fast rate. All 14 *Ss* in the non-replacement group reached criterion in the allotted hour, and generally in 15 to 30 min. Six of the 14 *Ss* in the replacement-group had not reached criterion at the end of the hour, and 60 trials were chosen as the cut-off point. By Kendall's test of significance of τ_{bis}, the difference in number of trials to criterion between the two groups was significant beyond the 2% level ($S = 134$, $SD_S = 43.5$).[5]

Since there were only enough different pairs of syllables to allow 15 completely different lists to be presented to *Ss* in the replacement-group (there was a total of 120 different pairs), some pairs were used a second time on the assumption that if *S* did not respond correctly to a pair, and then saw 119 different pairs after it, repetition was in effect a new pair. However, because of a plausible objection to this procedure, the two groups were compared on the first 11 trials; this is the number of trials it took the fastest *S* to reach criterion. The mean number of correct responses on Trials 1 through 11 for both conditions is shown in Fig. 1. Again using Kendall's test of significance of τ_{bis}, the replacement and non-replacement groups were compared on each trial from 1 through 11. The differences between the two groups are not significant for Trials 1 through 6 and are significant beyond the 5% level for Trials 7 through 11.

At the end of the first trial, the only identical trial for both groups, the total number of correct responses was 6 for the non-replacement group and 7 for the replacement group, from a total possible of 112 correct responses. After two trials, these figures were 6 and 10. Thus, there is no indication of an initial difference favoring the non-replacement group.

[5] M. G. Kendall, *Rank Correlation Methods*, 1948, 44.

Slow rate. In the replication of the one-trial study, the mean number of trials to criterion was 9.5 for the non-replacement group and 10.6 for the replacement-group. The *SD*s were 3.6 and 3.3. The difference between the two groups in number of trials to criterion was not significant ($S = 15$, $SDs = 24.8$, $P < 0.56$). Since essentially the same results were obtained as reported by Rock, only 10 *S*s were used in each condition.

Fast rate vs. slow rate. A comparison of the four groups in time to reach criterion shows that the *S*s in the non-replacement groups required

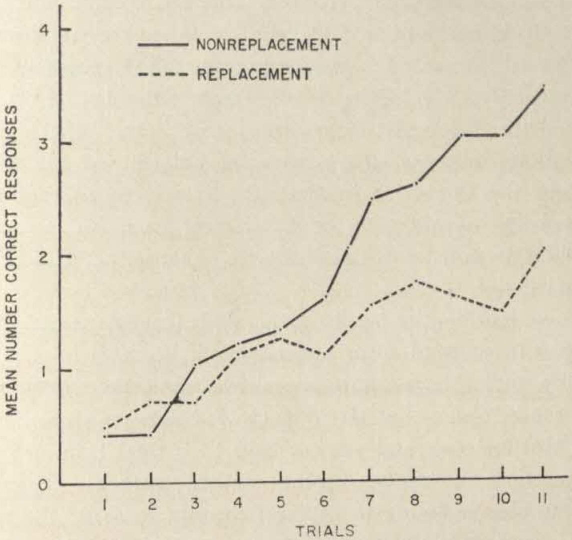


FIG. 1. MEAN NUMBER OF CORRECT RESPONSES ON TRIALS 1–11 FOR THE REPLACEMENT AND NON-REPLACEMENT GROUPS IN THE RAPID RATE OF PRESENTATION

a mean total time of 21.2 min. for the slow rate, and 24.0 min. for the fast rate of presentation. The replacement-group in the slow condition reached criterion in a mean total time of 23.7 min. Since six *S*s in the replacement-group of the fast condition did not reach criterion, the measure is meaningless for this group. If, however, those six *S*s are treated as if they had reached criterion on their last trial, the mean time is 42.5 min. for that group. These times include all rest periods.

DISCUSSION

One assumption of this study was that reducing all interval-times reduces rehearsals. This assumption seems to be supported by the fact that

the mean times to criterion, for the three groups which were allowed either repetition or rehearsal, were nearly identical. These times were 21.2 24.0 and 23.7 min. The one group which was allowed neither rehearsal nor repetition (the replacement-group with fast rate) required a mean of more than 42.5 min. to reach criterion. An interesting hypothesis which follows is that amount learned is a function of time, provided repetition or rehearsal is allowed.

The results indicate that some kind of repetition is required to learn associations which have been perceived. This can account for the results obtained by Clark, Lansford, and Dallenbach,[6] when they replicated Rock's study which used letter-number pairs[7] except that the exposure-time for the pairs was reduced to 1 sec. The *S*s were presented a list of material at a fast rate and then allowed time in which to rehearse whatever material was in short-term memory. If *S*s rehearsed and learned any pairs incorrectly during the inter-trial interval, the interference from those incorrect associations would give the replacement-group an advantage since they would be presented new pairs in place of those they had failed. Such are the results.

Since it appears from the results of this study that *S*s require opportunity for repetition in order to form associations, it becomes important not to ask if Rock's studies demonstrate one-trial learning, but to determine why his results can be obtained at all. Why can *S*s be shown material, be given time with that material, and yet not have their total time or number of trials to criterion affected by removal of that material? Three possible answers are: (1) *S*s decide to rehearse covertly some of the pairs until they have learned them and ignore, or not even look at, other pairs during this time; (2) *S*s intend to learn every pair, and look at every pair, but do not rehearse all of them; and this 'looking at only' does not affect learning; (3) *S*s look at, and actively rehearse, all of the pairs but learn nothing measureable about some of them.

If (1) is what is happening, determining why Rock's results can be obtained is uninteresting. From *E*'s observations this is not, however, the case. The *S*s appear to attend to all of the pairs. But it has been reported that when *S*s are forced to attend to each pair, by being required to respond overtly with each of the pairs as they are presented, Rock's results disappear even at the slow rate of presentation.[8] This result is not unex-

[6] Clark, Lansford, and Dallenbach, *op. cit.*, 36.
[7] Rock, *op. cit.*, 1957, 187.
[8] Leo Postman, Personal communication, 1960.

pected when one considers that the overt response lessens the available time for forming each association.

If (2) describes what is happening, it means that seeing without rehearsing does not alter the probability of *S*'s giving a correct response. The next possibility (3) is an extension of (2) and suggests that one can look at, and actively rehearse, some material without this effort's having any effect on learning. If either (2) or (3), or a combination of them, is correct further investigation is demanded.

SUMMARY

A study by Rock which appears to demonstrate that *S*s form associations on one trial rather than by gradual strengthening was replicated (1) as run by Rock, and (2) with rate of presentation increased. Results show that evidence for one-trial learning disappears when rehearsal-time available to *S*s is eliminated or greatly reduced.

THE EFFECT OF REPETITION ON THE RETENTION OF INDIVIDUAL WORDS

By BENNET B. MURDOCK, JR., and ARTHUR J. BABICK,
University of Vermont

Recently Rock has presented evidence suggesting that repetition prior to the formation of an associative bond does not facilitate learning,[1] and this finding has subsequently been confirmed in two different laboratories.[2] However, as Clark, Lansford, and Dallenbach have pointed out,[3] further corroboration with different learning methods and materials would appear to be highly desirable. Estes has reported data which appear to substantiate Rock's conclusion.[4] The present experiment is a further study of the effects of repetition on retention.

Procedure. We used the method of free recall. Lists of 25 words were read to *S* by *E;* after each list *S* recalled as many of the 25 words as possible. One of the 25 words was the 'critical word' (*CW*). If the *CW* was not recalled, it was repeated in the next list. If it was not recalled after the second presentation it was repeated in the third list, and so on. Thus, the *CW* reappeared in list after list until it was correctly recalled; then a new *CW* was introduced and the same procedure continued.

With the above procedure we can determine the effect of repetition on retention by comparing the recall of *CWs* with the recall of non-*CW*s (*i.e.* comparable words presented only once). If repetition does have an effect, then with each additional repetition the probability of recall of the *CW* should increase until it is correctly recalled. On the other hand, if repetition has no effect, the probabilities of recall of *CW*s and non-*CW*s should not differ.

Each *S* was given a total of 80 lists: 30 in the first session, 30 in the second session, and 20 in the third session. The stimulus-material consisted of words taken from the Thorndike-Lorge list of the 1000 most common words in the English language.[5] There were altogether 25 *CW*s. These 25 words were randomly selected

* Received for publication August 3, 1960. This study was supported by a research grant, M-3330, from the National Institutes of Health.

[1] Irvin Rock, The role of repetition in associative learning, this JOURNAL, 70, 1957, 186-193; Irvin Rock and Walter Heimer, Further evidence of one-trial associative learning, this JOURNAL, 72, 1959, 1-16.

[2] Michael Wogan and R. H. Waters, The role of repetition in learning, this JOURNAL, 72, 1959, 612-613; L. L. Clark, T. G. Lansford, and K. M. Dallenbach, Repetition and associative learning, this JOURNAL, 73, 1960, 22-40.

[3] Clark, Lansford, and Dallenbach, *op. cit.*, 38.

[4] W. K. Estes, Learning theory and the new "Mental Chemistry," *Psychol. Rev.*, 67, 1960, 207-223.

[5] E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944, 267-270.

from the 1000 words and were the same for all *S*s. They were (in order): simple, dollar, remain, president, agree, now, double, buy, going, number, lost, always, from, subject, became, follow, less, shape, come, upon, their, leave, might, well, and happen. For each list, the 24 other words were selected from the remaining 975 words. With 80 lists and 24 words per list a total of 1920 words was needed. Thus, almost all of the 975 words had to be used twice. The lists were constructed by arranging the 975 words in 40 rows of 24 each; the second set of 40 lists was obtained from the columns instead of the rows. No two words ever appeared in the same list more than once, and no word (other than a *CW*) was ever repeated in two successive lists.

Because of marked serial-position effects the *CW* was always placed at or near the middle of the list where the curve is flattest.[6] Specifically, the *CW* was randomly assigned to Serial Positions 10-16, inclusive. As far as possible, each of these positions was used equally often. There were six non-*CW*s per list; these were the remaining words in Serial Positions 10-16. Thus, if repetition has no effect, the average probabilities of recall of *CWs* and non-CWs should be the same.

There were 18 *S*s, students of both sexes from the course in introductory psychology. They were tested individually. As soon as a list had been read, *S* recalled as many words as possible in any order he wished. When 20 sec. had elapsed without any new words being recalled the trial was terminated and the next list read to *S*. Two practice-lists preceded the first session, and one practice-list preceded the second and third sessions. A 3-min. rest was given after every 10 test-lists. The *S*s were not told that one word would be repeated until it was recalled; if during the course of testing they asked (about one-quarter did) they were given an evasive answer.

The order in which the 80 lists were presented differed from *S* to *S*. The order in which the *CW*s were presented was the same for all *S*s. Thus, except for chance occurrences, each *CW* appeared in different lists for each *S*.

Results. Since the presentation of a new *CW* depended upon *S*'s performance, the number of *CW*s presented varied from *S* to *S*. The mean number of *CW*s presented was 11.1 words with an *SD* of 4.6.

For all *S*s, the total number of non-*CW*s recalled was 1,345 words. With a theoretical maximum of 8640 words (6 serial positions, 18 *S*s, and 80 lists per *S*) the over-all probability of recall of a non-*CW* was 0.156. To determine if the *CW*s were of comparable difficulty they were analyzed in the same way. For all *S*s a total of 199 *CW*s was presented at least once, and of these 32 were recalled after the first presentation. The over-all probability of recall of a *CW* after the first presentation was 0.161, and these two probabilities do not differ significantly ($z = 0.19$). Since the initial probabilities of recall of *CW*s and non-*CWs* were almost identical, it would appear that *CW*s and non-*CW*s were of comparable difficulty.

[6] James Deese and R. A. Kaufman, Serial effects in recall of unorganized and sequentially organized verbal material, *J. exp. Psychol.*, 54, 1957, 180-187; see Fig. 1, p. 182.

The basic hypothesis of the present experiment is that the repetition of a word does not affect the probability of its recall. If this hypothesis is correct, then (with p the probability of recall and q the probability of non-recall) the probability that a *CW* will be recalled after the first presentation is p, after the second presentation pq, after the third presentation pq^2, after the fourth presentation pq^3, etc. Thus, the hypothesis to be tested is that the obtained distribution of recall-scores for *CW*s can adequately be described by the geometric distribution pq^{x-1} where x is the ordinal number of the presentation (*i.e.* 1, 2, 3, . . .).

In testing the obtained distribution against the expected geometric distribution, all the data on recall of *CW*s were used. That is, those *CW*s

TABLE I

Expected (pq^{x-1}) and Obtained Number of Presentations (x) Required for Recall of *CW*s

x	Expected	Obtained	x	Expected	Obtained
1	31.0	32	9	8.0	5
2	26.3	25	10	6.8	8
3	22.1	28	11	5.8	5
4	18.7	18	12–13	8.8	8
5	15.7	16	14–15	6.2	4
6	13.3	9	16–18	6.2	6
7	11.1	12	19+	9.4	15
8	9.6	8			
			Total	199.0	199

that were recalled after the first presentation were included as well as those *CW*s that required one or more repetitions. When *S* did not recall the last *CW* presented to him, we assumed that he would have done so after the next presentation (*i.e.* the eighty-first trial). Such an assumption is patently false, but it works against the hypothesis under investigation and makes it possible to include 15 more recall-scores of *CW*s in the data. To set up the expected geometric distribution we used the value $p = 0.156$, the probability of recall of the non-*CW*s. The expected and obtained distributions are shown in Table I. As can be seen, the differences are rather small and non-systematic. The value of Chi-square was 9.07 which, with 13 *df.*, was far from significant. By chance alone a worse fit would be expected more than 75% of the time. Thus, it would appear that the obtained results can be described quite accurately by a geometric distribution, and the results show that repetition has little effect on the retention of individual words.

The *CW*s used in the present experiment were, not surprisingly, of unequal difficulty. An analysis of variance of the number of trials required

to recall the first six *CW*s (the only *CW*s recalled by all 18 *S*s) showed that the differences among the words were significant at well beyond the $\frac{1}{10}\%$ level. A difficult word is one with a low probability of recall while an easy word is one with a high probability of recall; thus, the *CW*s differed in their probability of recall. Fortunately it is possible to obtain a numerical estimate of the individual p-values for at least some of the *CW*s if one is willing to assume that these p-values do not differ from *S* to *S*. The p-value for any given *CW* is simply the reciprocal of the mean number of presentations required to recall it. This follows from the fact that, in a geometric distribution, the mean is equal to $1/p$. The p-values for the first 12 *CW*s were as follows: simple, 0.17; dollar, 0.31; remain, 0.07; president, 0.47; agree, 0.11; now, 0.12; double, 0.12; buy, 0.10; going, 0.13; number, 0.17; lost, 0.18; and always, 0.11. The number of *S*s recalling more than 12 *CW*s was too scanty to justify computing p-values for the remaining words.

The differences among *CW*s in p-values immediately raises a question about averaging. That is, in the Chi-square reported above, one p-value (0.156) was used to calculate the expected distribution, but the group-curve (that is, the expected distribution shown in Table I) may not accurately reflect the average of the individual curves based on different p-values for different *CW*s. This problem of averaging has been discussed by Estes,[7] and the geometric distribution used here falls into Class *C*, functions modified in form by averaging. In such a case, however, Estes points out that the distortions due to averaging will be a function of the dispersion of parameter-values (in the present case, p) and that sometimes this can be determined empirically.[8] In the present case we have the p-values for the first 12 *CW*s. We set up the appropriate geometric distribution for each of the different p-values obtained and then averaged; while the results were very close to the expected results shown in Table I (no differences in probability greater than 0.02), there was a noticeable curvature in a semi-log plot. The obtained results were tested by Chi-square against the results expected by averaging the different p-values; the fit was just as good as before ($p = 0.75$). Thus, the results are adequately described by a geometric distribution even when allowance is made for unequal difficulty of the several *CW*s.

In the preceding analysis it was necessary to assume that the p-values of the *CW*s did not differ from *S* to *S*. In view of the possibility of idio-

---

[7] Estes, The problem of inference from curves based on group data, *Psychol. Bull.*, 53, 1956, 134-140.

[8] Estes, *op. cit.*, 138.

syncratic factors such an assumption may not be tenable. Also, the two Chi-squares reported above may be questioned because of lack of independence (the 199 recall-scores were based on multiple entries from 18 *S*s). Therefore, still a third method of analysis was undertaken. We determined the probability of recall of non-*CW*s for each *S* individually; these *p*-values ranged from 0.106 to 0.235. Then, using the reciprocal relationship between *p* and mean trials to recall, we determined for each *S* the expected mean *CW*-recall. If a given *S* had a *p* of 0.20, then we would expect that the mean number of trials required for that *S* to recall the *CW*s would be 5.0 trials.

A comparison of the expected and obtained results (again assuming recall on the eighty-first trial) showed that 12 out of 18 *S*s required more trials than expected, while only 6 *S*s required fewer trials ($p > 0.05$). The difference in number of trials between observed and expected values was in the same direction (mean of 1.0 trials) but again this difference was not significant ($t = 1.68$, $p > 0.05$). Thus, if anything, *S*s required more trials to recall the *CW*s than one would expect on the basis of individual *p*-values, and there is no evidence that idiosyncratic factors affected the main results.

Discussion. Had repetition benefited retention, the recall-curves for the *CW*s would have risen further and further above the expected geometric curve as the number of repetitions increased. Irrespective of differences among *CW*s and possible idiosyncratic differences among *S*s, the results obtained in the present experiment consistently followed a geometric curve. Therefore, the probability that a *CW* would be recalled on any given trial was completely independent of the number of times that that particular *CW* had been presented before. Our results clearly substantiate Rock; repetition prior to the first correct recall did not benefit retention.

Could the results of the present experiment be explained (or explained away) on the grounds that any word not recalled on the first presentation by a given *S* must be more difficult for him, with this increased difficulty counteracting the benefits of repetition? We think not. Differences in difficulty are really immaterial. Such differences merely mean that the *p*-values of individual *CW*s differ; as has been pointed out, these differences complicate the problem of averaging but do not affect the main results. The real test of the effects of repetition is in producing departures from the geometric distribution, and there simply was no evidence of any such departure in the present results.

Finally, it should be noted that both Miller and McGill[9] and Bush and Mosteller[10] have assumed that repetition prior to the first correct recall is ineffective. In the Bush-Mosteller stochastic model this is represented by the assumption that $\alpha_2 = 1$, and the authors present some supporting evidence from free-recall verbal learning.[11] Actually, both of these stochastic models go even farther and assume that *any* non-recall leaves response-probability unchanged. According to both these models, however, response probability is increased by correct recalls; the recent results of Estes[11] appear to provide direct support for this contention.[12]

Summary. The present experiment used the method of free recall to study the effect of repetition on learning. One critical word was placed in the middle of a list of 25 words read once to *S*, and this same critical word was repeated in list after list until it had beeen correctly recalled. There was no apparent effect of repetition on recall.

[9] G. A. Miller and W. J. McGill, A statistical description of verbal learning, *Psychometrika*, 17, 1952, 369-396.

[10] R. R. Bush and Frederick Mosteller, *Stochastic Models for Learning*, 1955, 217-236.

[11] Bush and Mosteller, *op. cit.*, 218-223.

[12] Estes, *op. cit., Psychol. Rev.*, 67, 1960, 215-220.

THE EFFECT OF A REMOTE ANCHOR UPON JUDGMENT WITH A SALIENT WITHIN-SERIES STIMULUS-OBJECT AND WITH A FREE CHOICE OF SCALE

By SAMUEL FILLENBAUM, University of North Carolina

The judgment of stimuli is influenced by their context. It has been a very common finding that, upon the introduction of an anchor-stimulus outside one end of the stimulus-distribution, there are shifts in the judgment of the series-stimuli away from the category assigned to the anchor-stimulus as the subjective scale of judgment extends toward the anchor-stimulus.[1] Thus, if an item smaller than any in the stimulus-distribution is added, the near end of the judgmental scale shifts toward the anchor-stimulus, the category-limens shift downward, and there is an increase in the scale-values of the series-stimuli which now are placed in somewhat higher categories.

The primary purpose of this study was to see if these changes in judgment would still be obtained if the stimulus-series contained a salient member which 'naturally' defined one of the categories of the judgment scale—the middle category. The *O*s were required to rate on a slimness-broadness dimension a series of rectangles symmetrically spaced around the central rectangle—a square. It was thought that the introduction of an anchor-stimulus, a rectangle slimmer than any in the experimental series, would lead to the usual changes in the judgment of the rectangles on the side of the square toward the anchor—the rectangles slimmer than the square—but would not lead to any changes in the judgment of the square or the other rectangles broader than the square, on the far side of the distribution away from the anchor.

As has been pointed out before, the interpretation of context-effects in judgment is ambiguous.[2] They can be regarded "either as distortions of identity judgments which are independent of the specific details of the response system allowed, or as semantic effects limited to the specific

* Received for publication August 16, 1960.

[1] W. A. Hunt and John Volkmann, The anchoring of an affective scale, this JOURNAL, 49, 1937, 88-92; H. R. McGarvey, Anchoring effects in the absolute judgment of verbal materials, *Arch. Psychol.*, 39, 1943, (No. 281), 1-86.

[2] D. T. Campbell, N. A. Lewis and W. A. Hunt, Context effects with judgmental language that is absolute, extensive, and extra-experimentally anchored, *J. exp. Psychol.*, 55, 1958, 220-228.

response system employed in the method of single stimuli."[3] This study was designed, also, to get some information as to whether or not the judgmental changes resulting from the introduction of a remote anchor are mainly semantic effects. If *O* has to make his judgments using a fixed number of categories in the scale, the introduction of a remote anchor may lead to "a redefinition of the judgment categories which can be independent of any distortion in identity judgments."[4] The whole experiment was therefore run under fixed-scale instructions and repeated under instructions which permitted the *Os* to use as many categories as they wished in their subjective scales of judgment.

Experimental design. For some of the *Os*, the distribution of stimulus-objects consisted of rectangles both slimmer and broader than a square, and included a square as the perceptually salient item (Distinguished Condition). The other *Os* were presented with a homogeneous stimulus-distribution, all items being slimmer than a square (Undistinguished Condition). In each condition, some of the *Os* were permitted free choice as to the number of categories to be used in the judgmental scale and other *Os* were told to use a fixed five-category scale. Finally, in each of the four sub-conditions some of the *Os* were presented with the control distribution in which all the stimulus-objects were evenly spaced over the range, and other *Os* were presented with the experimental distribution which was the same as the control distribution with the addition of a remote anchor two steps below the slimmest rectangle in the series. The anchor was presented for judgment just as were the other objects, but it appeared twice as often as any other one.

Materials. Each rectangle was displayed on a 31 × 56 cm. white card in front of a large dark screen for about 2 sec. The height of the rectangles was kept constant, and shape was varied by changing the breadth.

In the Distinguished Condition, the seven rectangles were spaced at equal logarithmic steps. The ratio of height to base for the slimmest rectangle was 1.21:1, while for the broadest rectangle this ratio was 0.73:1. The height of the rectangles always was 20 cm., and breadths were as follows: 16.50, 17.54, 18.68, 20.00, 21.26, 22.72, and 24.08 cm. The ratio of height to base for the anchor was 1.38:1, and its breadth was 14.50 cm.

In the Undistinguished Condition, the six rectangles employed also were spaced at equal logarithmic steps. The ratio of height to base for the slimmest rectangle was 5.02:1, while for the broadest rectangle this ratio was 2.64:1. The height of the rectangles always was 25 cm., and their breadths were as follows: 4.98, 5.65, 6.45, 7.33, 8.33, and 9.47 cm. The ratio of height to base for the anchor was 6.44:1, and its breadth was 3.87 cm.

Instructions and procedure. The *Os* were give sheets on which to record their responses and then read the following instructions:

This is an experiment in shape-perception, in the perception of the shape of rectangles. You will be shown a series of rectangles. As each rectangle is presented, make a judgment of how slim or broad it seems to you. [For the fixed-

[3] *Ibid.*, 220.
[4] *Ibid.*, 221.

scale *Os*, the instructions continued thus]: Use the following five-category scale: −2 (very slim), −1 (slim), 0 (medium), +1 (broad), +2 (very broad). [For the free choice *Os*, the following was substituted]: Use as many categories as you can in your scale of slimness-broadness. Let 0 stand for medium and use +1, +2, etc., as you may need for the broader rectangles and −1, −2, etc. as you may need for the slimmer ones, so that your largest positive number will be used for the broadest rectangle and your largest negative number for the slimmest rectangle. [*The instructions up to this point were repeated for all* Os, *and then continued as follows*]: Look at the rectangle, make your judgment and record it on the sheet. Make a judgment for every rectangle even though the decision may sometimes be difficult. Do not go back and change your response after new rectangles have been presented. This experiment deals with subjective impressions, and since you are to judge how slim or broad a rectangle looks to you on a particular presentation there cannot be any right or wrong answers. Before you actually start judging you will be given some practice just looking at the rectangles one at a time, so as to acquaint yourself with them.

In the Distinguished Condition, rectangles were randomized in blocks of seven (control distribution), or nine (experimental distribution). In the Undistinguished Condition, rectangles were randomized in blocks of six (control distribution), or eight (experimental distribution). The acquaintance-series comprised from 24-28 presentations (3-4 blocks); the *Os* then made 80 judgments. The last 60 or so of these constituted the data, the first few judgments not being scored to allow further time for the judgmental scale to stabilize.

Observers. The *Os* were 205 undergraduates at the University of California who participated in the experiment to satisfy a class requirement.[5] They were run in small groups, mostly of three or four, and were assigned at random to the various conditions: *Distinguished Condition,* fixed scale, 22 experimental *Os* and 18 control *Os*; free scale, 27 experimental and 25 control; *Undistinguished Condition,* fixed scale, 21 experimental and 43 control; free scale, 24 experimental and 25 control.

Results. For all *Os* using the fixed five-category scale, the mean rating given each rectangle was computed. The means of these mean ratings for the various conditions are given in Table I, as are the *t*-values resulting from an analysis of the differences in rating between the experimental and control *Os*. It is clear that in these circumstances the introduction of a remote anchor led to significant shifts in the judgment of the rectangles both in the Distinguished and in the Undistinguished Condition.

For those *Os* who were permitted free choice as to the number of categories in the judgmental scale, the data were pooled by noting for each *O* the location of his neutral level. This was done by determining the rectangle nearest which the 'medium' judgment fell. If the remote anchor was effective, then the scale should extend toward it and the neutral level should be pulled down. For both the Distinguished and the Undistinguished Condition, the distribution of the neutral level was determined for

[5] For various reasons the data of 18 or 19 *Ss* had to be discarded; most of these *Ss* did not follow the instructions or turned in apparently random response-sheets.

the control and for the experimental *O*s. The Mann-Whitney test was applied to these values as a test of the difference in average location of the neutral level. For both conditions, the neutral level was lower in the experimental than in the control group, but the significance of these differences is borderline (in the Distinguished Condition, $p = 0.06$, in the Undistinguished Condition, $p = 0.04$, both tests one tailed).

The results were essentially the same whether the distribution of the stimulus-rectangles was homogeneous or contained a perceptually salient member which 'naturally' defined the middle category (for the control

TABLE I

MEAN CATEGORY-NUMBER ASSIGNED EACH STIMULUS-RECTANGLE AND *t*-VALUES FOR SIGNIFICANCE OF DIFFERENCES IN RATING BETWEEN EXPERIMENTAL AND CONTROL GROUPS

Condition	Group	Anchor	Stimulus-number						
			1	2	3	4	5	6	7
Distin-	Exp.	−1.98	−0.91	−0.69	−0.25	0.28	0.92	1.44	1.96
guished	Control		−1.86	−1.12	−0.57	0.02	0.61	1.15	1.89
	t		11.40*	3.59*	2.15†	2.17†	2.48*	2.62*	0.93
Undistin-	Exp.	−1.74	−0.73	−0.27	0.27	0.79	1.26	1.81	
guished	Control		−1.71	−1.02	−0.38	0.32	0.92	1.68	
	t		11.12*	7.02*	5.48*	4.42*	3.35*	1.29	

* $p<0.01$ (one tail).
† $p<0.05$ (one tail).

group the mean category-number assigned the square was 0.02). The range in width of the rectangles presented had a decided effect on the rating of the items when a fixed scale was specified. Judgment was not 'absolute' when the distribution of rectangles extended farther toward the narrow than the broad end, a square was rated somewhat broader than medium on a scale of slimness-broadness.

When free choice was permitted as to the number of categories in the subjective scale, the differences in judgment due to the presence of a remote anchor, while still in the same direction, were less clear cut, both for the homogeneous and the heterogeneous stimulus-distributions. This finding indicates that, in part, these differences may perhaps be regarded as semantic effects resulting from restriction in the number of available response-categories. The finding that even in the free-choice case there were still differences of borderline significance suggests, however, that these context-effects cannot be entirely explained by recourse to semantic factors, and that shifts in stimulus-range may lead to changes in the appearance of the stimulus-objects which are independent of the particular language employed by *O*.

The addition of another rectangle and the broadening of the stimulus-range for the experimental groups did not lead to any increase in the number of categories employed when free choice of scale was permitted. Splitting the distribution of the number of categories employed at the median of the total group (experimental and control *O*s) resulted, for both conditions, in X^2- values far short of significance.

Summary. This study was designed to determine if the changes in judgment usually produced by a remote anchor would occur when the stimulus-series contained a perceptually salient member which 'naturally' defined the middle category of the judgmental scale. The study also was designed to provide some information as to whether these changes in judgment might be regarded mainly as semantic effects due to restriction in the number of available response-categories.

The *O*s were required to rate each of a series of rectangles on a slimness-broadness dimension. The results were much the same whether the distribution was homogeneous (all rectangles slimmer than a square) or heterogeneous (including a square as the perceptually salient member and rectangles both slimmer and broader than the square). The presence of a remote anchor led to differences in judgment, the rating of the stimulus-series shifting away from the category assigned the anchor. This effect was quite clear cut when a fixed five-category scale was employed, but only of borderline significance when free choice was permitted as to the number of categories in the scale.

The results indicate that the range of stimulus-objects employed has an effect on judgment. Even with a salient anchor within the stimulus-distribution, judgment was not 'absolute.' The fact that, with free choice of categories, there were still some changes due to the presence of the remote anchor suggests that this effect cannot be interpreted entirely in semantic terms.

ATTENTION AND ASSIMILATION

By RILEY W. GARDNER and LEANDER J. LOHRENZ,
The Menninger Foundation

In a recent study of individual consistencies in cognitive functions, Gardner and his associates[1] stated that consistent individual differences in assimilation,[2] which they observed and referred to as a 'leveling-sharpening dimension,'[3] could result from differences in the degree, *i.e.* clarity and vividness, with which persons attend to new experiences. This hypothesis is related to Köhler and Adams' finding that perceptual articulation and degree of satiation are functions of the degree of attention.[4] Clarification of relations between attention and assimilation is important to understanding both the phenomenon of assimilation and the wide ranges of individual differences observed in studies of assimilation.

The double task method, devised by Geissler,[5] was employed to reduce the degree of attention given to visual designs by an experimental group, as compared to a control group who viewed the designs under normal conditions. There seemed to be no doubt that requiring *Ss* in the experimental group to count backward by twos while viewing designs presented sequentially reduced attention to the designs. All members of a pilot group reported that counting backward was surprisingly distracting, that it interfered with attending to the designs and made it more difficult to recall them.

* Accepted for publication September 14, 1960. This research was supported by Research Grant M-2454 from the National Institute of Mental Health, Public Health Service.

[1] R. W. Gardner, P. S. Holzman, G. S. Klein, H. B. Linton, and D. P. Spence, Cognitive control: A study of individual consistencies in cognitive behavior, *Psychol. Issues*, 4, 1959, 138.

[2] Wolfgang Köhler, Zur Theorie des Sukzessivvergleichs und der Zeitfehler, *Psychol. Forsch.*, 4, 1923, 115-175; Otto Lauenstein, Ansatz zu einer physiologischen Theorie des Vergleichs und der Zeitfehler, *ibid.*, 17, 1933, 130-177.

[3] P. S. Holzman, The relation of assimilation tendencies in visual, auditory, and kinesthetic time-error to cognitive attitudes of leveling and sharpening, *J. Pers.*, 22, 1954, 375-394; P. S. Holzman and G. S. Klein, Cognitive system-principles of leveling and sharpening: Individual differences in assimilation effects in visual time-error, *J. Psychol.*, 37, 1954, 105-122.

[4] Wolfgang Köhler and P. A. Adams, Perception and attention, this JOURNAL, 71, 1958, 489-503.

[5] L. R. Geissler, The measurement of attention, this JOURNAL, 20, 1909, 515. See also K. M. Dallenbach, The measurement of attention, *ibid.*, 24, 1913, 490.

METHOD

Subjects. The *Ss*, 30 women, ranging in ages from 22 to 38 yr. (mean 30.1 yr.), were divided haphazardly into two groups of 15 each: one, the control group, and the other, the experimental group.

Apparatus. Four pairs of designs (See Fig. 1) conducive to mutual assimilation were selected from those employed by Gibson.[6] Each design was displayed individually on a white screen by means of an automatic slide projector connected to an interval-timer. Room lights were dimmed to increase the visibility of the designs.

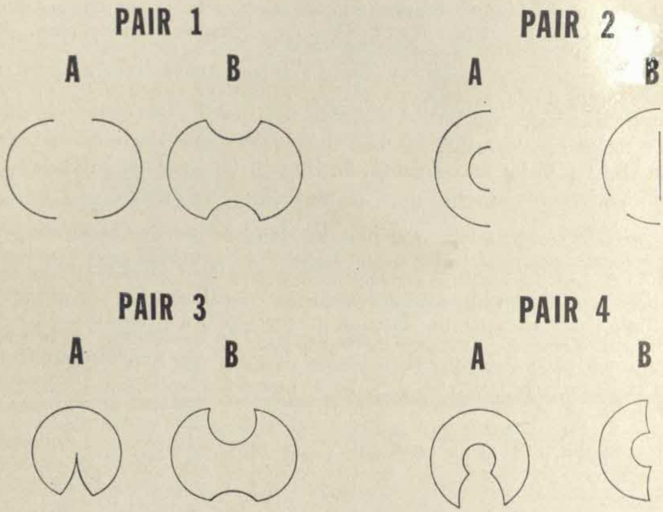


FIG. 1. DESIGNS EXPOSED SUCCESSIVELY TO THE *Ss*

The first design of each pair (*e.g.* 1 A) was shown for 0.5 sec. After a 2-sec., blank interval, the second design of that pair (*e.g.* 1 B) was shown for 0.5 sec. The figures on the screen averaged about 32 in. in the longest dimension and were viewed at a distance of about 15 ft.

Instructions and procedure: Control Group. The instructions and procedure for the Control Group were as follows.

> You will be using the group of papers numbered one through four. You will be shown some figures on the screen. They will appear in rapid succession, so you will need to watch the screen carefully. Watch the screen very carefully so that you will see the first figures.

Design 1 A was then projected, followed by 1 B. *E* then said:

> Use Sheet No. 1 to draw the figures that you just saw. Draw as accurately as possible. Draw and number the figures in the order in which they appeared.

[6] J. J. Gibson, The reproduction of visually perceived forms, *J. exp. Psychol.*, 12, 1929, 24.

Then the designs of Pairs 2, 3, and 4 were shown in turn after similar instructions were given. In every instance the *Ss* were given as much time as necessary to reproduce the designs. A rest-interval, equivalent to the time required for the special instructions given to the Experimental Group, intervened between the presentation of Pairs 2 and 3.

Experimental Group. For Pairs 1 and 2, the procedure was the same as that used with the Control Group. Before presenting the designs of Pairs 3 and 4, *E* said:

> We will now practice doing something that will be necessary for our next test. We will be counting backward by twos in unison. Since all of you will participate, we will have to work out the proper timing together now. Just for practice, let's try this now, starting at 30. Watch the screen while you are counting. Ready? Begin. 30 . . . 28 . . . 26 . . . (etc.).

E engaged the timer at the count of 20. While the *Ss* continued counting, two blank projections appeared on the screen in sequence in the temporal arrangement used for the designs. During this and subsequent counting periods, a second *E* observed the *Ss*. Those who counted softly (as if to reduce the distraction) were urged to count in a "loud, clear voice." *E* then said:

> Now we are ready for the next test. You will be counting backward by twos just as we have practiced it, but sometime during the time you are counting, two designs [3 A and 3 B] will be shown on the screen in succession, just as the two light patches appeared while we were practicing. Therefore, while you are counting, you will also have to watch the screen at all times so that you will not miss seeing these designs. Remember to watch the screen carefully as we count. When the designs come on, *keep counting.* Do *not* stop counting until you are specifically instructed to do so, which will be after the designs disappear. We will start with 60. Ready? Begin: 60 . . . 58 . . . 56 . . . (etc.).

E engaged the timer at the count of 52. He interrupted with the following instructions after both designs had been exposed and two additional numbers had been called.

> Stop counting. On Sheet No. 3 draw the figures you just saw as accurately as possible. Draw and number them in the order in which they appeared.

A similar procedure was used for Pair 4. Counting began at 80 and the timer was engaged at the count of 74. *Ss* again continued counting while viewing the designs.

Ratings of assimilation. Three raters independently ranked each of the four pairs of drawings by all of the 30 *Ss* in terms of degree of assimilative interaction. The lowest rank was assigned to the drawings containing least evidence of assimilation. The following general rules were formulated to help the raters distinguish assimilation from other kinds of inaccuracy in *Ss*' drawings.

(1) Assimilation was inferred only when one or both members of a pair varied from the actual stimuli in ways attributable to the presence of the other member of the pair or preceding pairs. Other variations from the actual stimuli were ignored.

(2) Omissions were not attributed to assimilation. When one member of a pair was not drawn, the rater based his judgment only on assimilation apparent in the other member.

(3) Between-pair interaction was considered more extreme evidence of assimilation than interaction between members of a pair.

Coefficients of concordance (Kendall's W) were computed for the rankings of each pair by the three raters. All four W's were significant at $P < 0.001$. The lowest of the four highly significant F-values was obtained for Pair 1, in which no interaction between the pairs was possible.

Results

Quantitative findings. Mann-Whitney U-tests were employed to assess differences between rankings for the control and experimental groups on Pairs 1 and 2, which both groups observed without counting, and on Pairs 3 and 4, which the experimental group viewed while counting back-

TABLE I
MANN-WHITNEY U-TESTS OF RATINGS OF ASSIMILATION FOR STIMULUS-PAIRS 3 AND 4

Rater	Sum of ranks		U	P
	control	experimental		
1	134.5	330.5	14.5	<0.002
2	135.5	329.5	15.5	<0.002
3	132.5	332.5	12.5	<0.002
Combined	129.0	336.0	9.0	<0.002

ward. None of the Us for the individual or combined raters was significant for Pairs 1 and 2. The sums of ranks, U- and P-values for Pairs 3 and 4 appear in Table I. The Experimental Group (condition of distraction) clearly showed greater assimilation than the Control Group (neutral condition).

Qualitative findings. Assimilation was clearly apparent in nearly all the drawings produced by our *S*s. We did not, however, anticipate the variety of forms in which assimilation occurred, particularly in the Experimental Group. These forms ranged from changes toward greater similarity between members of a single pair, or transposition of one element from one member to another, to gross interactions between pairs. Perhaps the outstanding example of the latter form of assimilation was produced by a member of the Experimental Group who, when asked to draw Pair 4, produced a recognizable drawing of Pair 1. Other forms of assimilation included rotations or partial rotations of one member of a pair into greater congruence with the orientation of the other member, or members of another pair. In an extreme example of such rotation, an *S* produced an unusual pair of drawings that appeared to contain elements of a preceding pair, but with the new pair rotated 90° in a sort of compromise with the earlier pair.

Discussion

Our results seem to provide clear evidence that the degree with which stimuli are attended to is a determinant of the amount of mutual assimilation among them when they are presented sequentially. Our results are thus analogous to Köhler and Adams' finding that attention is one determinant of perceptual articulation and degree of satiation.[7] Concentration of attention upon stimuli so fixes or stabilizes percepts that they are less susceptible to interaction with memories of related earlier experiences in the course of memory formation.

These results seem to imply that any artificial reduction of the degree with which *S* attends to a stimulus may increase the degree of assimilation it undergoes. There seems to be an obvious link, for example, between the present findings and the registration of stimuli presented tachistoscopically or in other ways that prevent optimal attention to the stimuli (*e.g.* as in masking procedures). Recent studies of subliminal registration are a case in point. Our results suggest that such stimuli may undergo a fate in the process of memory formation different from that of stimuli attended to with a greater degree of attention. In other situations, a low degree of attention to particular stimuli could result from the fact that *S*'s attention is primarily directed to other stimuli. To take but one example, our results suggest that impressions of 'incidental' stimuli in experiments on learning may be subject to greater assimilative interaction with related memories than are the stimuli to which attention is primarily directed.

From the point of view of individual differences in assimilation, the present results point to an even more intriguing possibility: that individuals differ in the discreteness with which they record and hence remember events in part because of characteristic differences in the degree with which they attend to stimuli.

[7] Köhler and Adams, *op cit.*, 503.

APPARATUS

THE RESISTIVE SHEET: A GRIDLESS AND WIRELESS SHOCKING TECHNIQUE

By NICHOLAS LONGO, LAWRENCE R. HOLLAND, and M. E. BITTERMAN, Bryn Mawr College

The difficulties inherent in the conventional shocking grid are well known. If the distance between bars is large, the animal may avoid contact with oppositely charged surfaces. If the distance between bars is small, the likelihood of short-circuiting by urine or faeces is high. In work with some species, such as the rat, the grid may be used with fair confidence if supplemented with a motor-driven commutator which produces a shifting pattern of potentials on the bars,[1] but a dependable grid is entirely out of the question with other species, such as the cockroach, whose extremities are rather delicate and whose capacity for the production of grid-shorting fluids is abundant.

As an alternative to the grid, wires may be attached directly to the animal; in the simplest case, only one lead is necessary because a conductive substrate will serve as the second. This technique works well with the pigeon, but it is inapplicable to many other animals. The rat, for example, shows great ingenuity in removing leads attached to it, and the problem of how to make a stable, conductive connection to the fly or the cockroach has thus far seemed insoluble.

Yet a third technique of administering shock, which involves neither grids nor wires, was explored several years ago by Licklider.[2] The animal stands on a single electrode which is connected to the output of a spark-coil, and the circuit is completed through the capacitance of animal to the walls of the experimental enclosure, which are grounded. Essentially the same method had been used two decades earlier by Griffith, who was dissatisfied, not with the shocking properties of grids, but with their tactual and visual properties.[3] Griffith wanted a shocking surface

* This work was supported by the Office of Naval Research under terms of Contract Nonr 2829(01). Reproduction of this report in whole or in part is permitted for any purpose of the United States Government.

[1] B. F. Skinner and S. L. Campbell, An automatic shocking-grid apparatus for continuous use, *J. comp. physiol. Psychol.*, 40, 1947, 305-307; L. B. Wyckoff and H. A. Page, A grid for administering shock, this JOURNAL, 67, 1954, 154; J. A. Dinsmoor. A new shock grid for rats, *J. exp. Anal. Behav.*, 1, 1958, 264.

[2] J. C. R. Licklider, A gridless, wireless rat-shocker, *J. comp. physiol. Psychol.*, 44, 1951, 334-337.

[3] C. R. Griffith, A new method of administering shock in animal experimentation, this JOURNAL, 43, 1931, 286-287.

which would be indistinguishable from the ordinary pathways of an elevated maze. Accordingly, he stapled a wire to a wooden pathway, covered the junction and the pathway itself with aluminum paint, and then connected the other end of the wire to one side of the secondary of a Harvard inductorium, the other side of the secondary remaining open. A rat standing on any part of the aluminum-painted surface would be shocked when a 12-v. source was connected across the primary, Griffith reported. One of the present authors (MEB) worked for several days not long ago with William A. Shaw of the University of Pennsylvania in an effort to duplicate Griffith's results, but sporadic shock was the best that could be achieved. A number of investigators attempted independently to duplicate Licklider's results soon after they were published, also without success. The report of Beck, Waterhouse, and Runyon well reflects the difficulties encountered.[4]

The technique which is described here—also gridless and wireless—grew out of our attempts to shock insects (flies and cockroaches). After many failures with grids of one sort or another, we succeeded at last by impressing a voltage across a piece of wet blotting paper on which the animal stood, thus connecting the animal in parallel with the circuit through the paper. (Leads were stapled to the paper and the junctions coated with silver circuit-paint.) Later, the wet blotting paper was replaced by a dry, hard-surfaced paper which was rubbed with powdered graphite or painted with a mixture of graphite and black asphaltum. In all cases, the electrical resistance of the shocking surface was so low to begin with (on the order of 100-500 Ω/in.) that it was not reduced appreciably by watery exudates from the cockroach. The shocking voltages used also were low (110-v. A.C. stepped down by means of a Variac). With this technique, failure of shock never was encountered in many hundreds of trials.

The method used to provide an objective measure of the efficacy of the new shocking technique (which was designed ultimately for the study of classical conditioning in insects) is an extension of a method designed earlier for the study of classical conditioning in small aquatic animals.[5] The transducer is a crystal phonograph-cartridge, the output of which is amplified and integrated. In this case, the cartridge detects, not the movement of a paddle in the water of an aquarium, but the movement of a platform on which the insect stands. An arrangement used for the fly (*Sarcophaga bullata*) is shown in Fig. 1. The platform is of thin plastic, painted with a mixture of graphite and black asphaltum, and flexible leads are run from each side of the platform to the output of the Variac which

[4] L. H. Beck, I. K. Waterhouse, and R. P. Runyon, Practical and theoretical solutions to difficulties in using Licklider's rat shocker, *J. comp. physiol. Psychol.*, 46, 1953, 407-410.

[5] J. L. Horner, Nicholas Longo, and M. E. Bitterman, A classical conditioning technique for small aquatic animals, this JOURNAL, 73, 1960, 623-626.

supplies the shocking voltage. The fly is restrained by an overhead beam running through a loop of thread fixed with wax to its dorsal surface. (The loop is attached in such a way that normal movement, even flight, continues to be possible when there is no restraint). The platform is set on a light rod which is inserted into the needle-holder of the phonograph-cartridge. In work with the cockroach (*Periplaneta americana* or *orientalis*), a covered box of thin plastic replaces the open platform used for the fly, and the unharnessed *S* is permitted to walk freely about in it. In both cases, the activity of the animal is measured during a 2-sec. interval (set

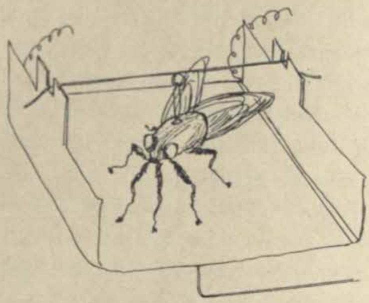


FIG. 1. AN ARRANGEMENT FOR MEASURING RESPONSE TO SHOCK IN THE FLY

on a Hunter timer). At the start of the interval, the cartridge is connected to the integrator, and at the termination of the interval the connection is broken. Shock may or may not be scheduled during the first portion of the interval, and the two kinds of interval are programmed in random order. The efficacy of the shocking technique may be evaluated by comparing the distribution of activity-scores (integrator-readings) for intervals scheduled to begin with shock and the distribution of activity-scores for control intervals. If there is no overlapping of the two distributions, one can say with considerable confidence that there has been no failure of shock, although overlapping does not mean necessarily that there has been failure.

Sample data for a cockroach (*orientalis*) and a fly are plotted in Fig. 2. For the fly, 80 v. were applied across a graphite-asphaltum surface; for the cockroach, 50 v. across a surface of moistened blotting paper. The curves indicate some variability in response to shock, but there is no suggestion of shock-failure; the measures for shock-intervals and control intervals do not overlap.

To adapt this technique for work with the rat, a shocking surface of higher resistance was required, because of the high skin-resistance of the rat. The higher resistance was achieved simply by using less graphite per volume of asphaltum in the graphite-asphaltum mixture, which was applied in this case to a base of Plexiglas. With the surface-resistance increased, a higher shocking voltage was required, and in the design of the shocking circuit one of the major difficulties inherent in the new technique was considered and solved. When a voltage-difference is applied at two points of a resistive sheet on which a rat stands, the amount of current

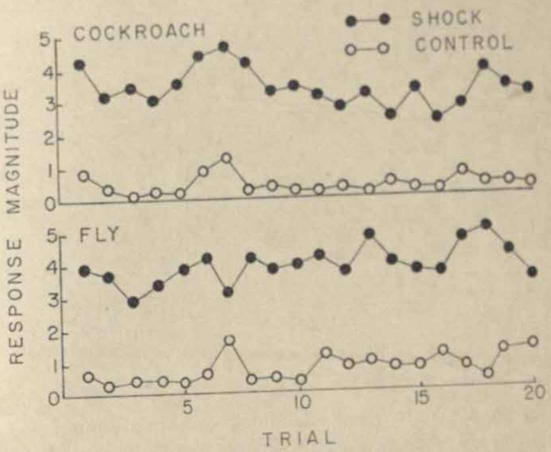


Fig. 2. Activity-Measures for Two Insects Made During Shock-Intervals and During Control Intervals

received by the animal will vary with its orientation. When the long axis of the animal is perpendicular to a line between the points at which the potential-difference is applied, it receives less shock than when its long axis is parallel to that line. To minimize the effect of orientation, a rotating electrical field may be used.

With 60-cycle A.C.-power, a rotating field is obtained by phasing two independent fields which are at right angles to each other in such a way that one reaches its peak value when the other is at zero. The rotation may be accomplished by either of two methods: (1) One field is retarded by a choke while the other is advanced by a condenser, with the sum of the phase-shifts being 90°. (2) One field is advanced almost 90° by a condenser while the other is equalized but not phased by a resistor. The second method is more economical in that the choke is eliminated, but a

higher supply-voltage is required for any given effect. The circuit actually used is diagrammed in Fig. 3, and the way in which it was connected with the shocking surface is shown. Two connections were made to each edge of the surface to promote a more uniform distribution of current. A more satisfactory (but more expensive) circuit of the first type replaces each R_1 with a choke of 200H; each capacitance is 0.04 μf., and each R_1 is 150K.

The shocking surface used in preliminary tests was a 9-in. square of ¼-in. Plexiglas, with a series of ¼-in. holes cut into it to permit the

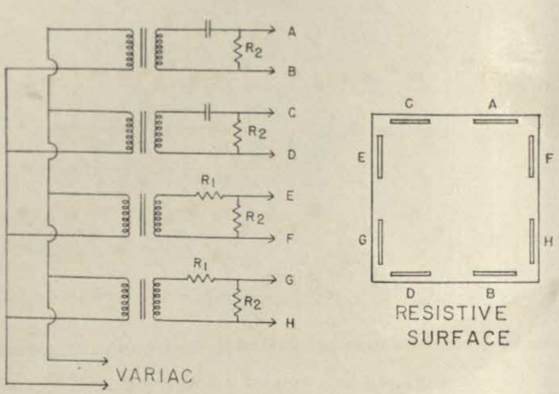


FIG. 3. A CIRCUIT USED FOR SHOCKING RATS AND PIGEONS
(Transformers: Stancor P-8150, 1550 v. Capacitors: 0.0025 μf.
Resistors: R_1, 1 megohm; R_2, 250K.)

running off of urine. At eight points equidistant about the edges of the square, banana plugs were set into the Plexiglas, and their junctions with the upper surface of the square were coated with silver circuit-paint. (The receptacles for the plugs were set into a base made of another piece of ¼-in. Plexiglas, and from them ran leads to the shocking circuit. This arrangement made it easy to remove the upper portion for cleaning.) A graphite-asphaltum mixture then was applied, which produced a resistance of about one megohm between opposite edges.

The base into which the shocking surface could be plugged rested on the floor of a small picnic chest. Below the base, fixed also on the floor of the chest, was a crystal phonograph-cartridge in a small metal case. From the needle-holder of the cartridge there protruded a light metal rod, to the other end of which was attached a lead weight. The picnic chest itself rested on a sheet of foam rubber. The arrangement was such

that slight movements of the rat on the shocking surface would displace the lead weight and produce a voltage across the cartridge. This voltage was amplified and integrated, in the manner described earlier, to provide an objective measure of overt response to shock.[6] (This apparatus, like that for insects, was designed ultimately for the study of classical conditioning.)

Some sample data for a rat are presented in the lower portion of Fig. 4. As in the work with insects, activity was measured during 2-sec. intervals, in the first second of which shock was scheduled. Activity also was measured during control intervals of the same duration, but the resulting values were too small to be worth plotting; the highest control value for the series of tests shown in Fig. 4 was 0.30. The Variac setting for the series was 50 v. The upper portion of Fig. 4 shows the results obtained with a pigeon. The procedure here was the same as that for the rat in all respects except that the pigeon's feet were rubbed with electrode-jelly before the tests were made. The highest control value recorded during the series was 0.03 (the pigeon tending, of course, to be rather inactive in the dark).

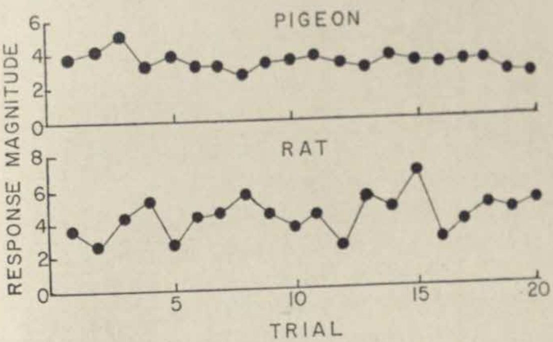


FIG. 4. RESPONSE TO SHOCK IN RAT AND PIGEON

While a painted shocking surface gives rather good results on the whole, it has certain disadvantages. Urine wets it easily, and markedly reduces its resistance in variable ways. Furthermore, after repeated exposure to waste-materials, and after repeated washing and drying, its homogeneity may be lost permanently. With somewhat greater effort, it is possible to make a more satisfactory surface and one which will last indefinitely. Rows of holes, 0.12-in. in diameter and 0.30-in. apart (center-to-center), are

[6] *Ibid.*, 623-626.

drilled in a square of $\frac{1}{4}$-in. Plexiglas, and filled with brass rivets (7/32-in. shanks and heads). Rows of larger holes, 0.20-in. in diameter, are drilled in staggered fashion between the rows of rivets to permit the running off of urine. To the underside of the plastic, a mixture of graphite and Weldwood Presto-Set Glue (in proportions adjusted to give the required resistivity) is applied with a trowel, the mixture being pressed into the hollow shanks of the rivets in such a way as to ensure good contact. When the glue has dried, a layer of clear plastic is sprayed over it, and the result is wet-proof.[7] It is the rivet-heads, of course, which constitute the shocking surface. Urine sprinkled on such a surface does not appreciably alter its resistance.

After rivet-head surfaces were constructed and found to work nicely for pigeon and rat, a much simpler solution to the problem of surface-stability was discovered. Enough carbon to give the desired resistance was added to a thick solution of Styrofoam in methyl-ethyl-keytone, and the mixture applied with a brush to a Plexiglas plate in the same manner as the carbon-asphaltum mixture used earlier. Again, holes were drilled in the plate to permit the running off of urine. Such surfaces have proved to be very stable indeed, and they are, of course, much easier to prepare than even the most primitive grid of equal dimensions.

A difficulty which has not yet been resolved, and which seems to be inherent in the technique here described, is to be found by the fact that the amount of shock received by the animal will vary with the nature of its contact with the shocking surface. All other things being equal, a rat receives less shock when it rears up on its hind legs than when it stands on all fours, and a pigeon receives less shock when it stands on one foot than when it stands on both feet. Just how serious this limitation is considered to be will depend on the purposes of the investigator. In work with some species, as has already been noted, there simply is no reasonable alternative to the resistive sheet. In work with other species, which permit a choice between grid and resistive sheet, the advantages and disadvantages of each method must be weighed. It is by no means certain, of course, that the intensity of shock from the best available grid is substantially less variable (quite apart from the possibility of complete failure) than the intensity of shock from a resistive sheet. It would be interesting to make comparative studies of the two techniques by methods such as we have used here.

---

[7] Glue rather than asphaltum is used as a base for the graphite because the plastic spray destroys the conductivity of the graphite-asphaltum mixture.

THE CONTINUOUS MEASUREMENT OF PALMAR SWEATING

By R. C. Wilcott, Western Reserve University, and H. G. Beenken, University of Nebraska College of Medicine

The development of an instrument for the continuous measurement of sweating has presented formidable problems. The method usually employed has involved the passage of dried air across the skin surface and the measurement of the moisture-content of the air as it leaves the skin. The major problem has been in finding a device that would provide a sensitive and reliable measure of the moisture-content of a continuous flow of air. The hydroscopic fibers employed by Darrow were satisfactory in this respect, but difficult to calibrate.[1] A device has recently become commercially available which overcomes this difficulty, and its application for the measurement of palmar sweating is described in this report.

The device used for the measurement of the moisture-content of a continuous flow of gas is a 'Coulometric Cell' developed and produced by Consolidated Electrodynamic Corporation (Pasadena, California). The following is a description of this unit based on the manufacturer's specifications: It operates on the basis of the simultaneous absorption and electrolysis of the water-content of a gas flowing through it. The cell consists of a tube which has a pair of platinum wires wound in a double helix on its inner surface. The space between the wires is coated with a film of hydrated phosphorous pentoxide. With a D.C.-potential applied to each of the platinum wires, the electrolysis-current is a linear function of the water-content of the gas passing through the cell. If gas-pressure, rate of flow, and temperature are held constant, the current through the cell will be proportional only to the water-concentration of the gas, with measurements from different gas-samples being directly comparable. Under these conditions, the absolute water-content of the gas can be expressed in parts per million by means of a calibration formula derived from Faraday's Law. The range of the cell is from 1-20,000 parts per million. The latency of the cell is 0.7 sec.

The accessory equipment used for adapting the cell to this problem

* This study was supported by NIMH Grant No. M-2603 (A) and by research funds of the Board of Control of the State of Nebraska.

[1] C. W. Darrow, Quantitative records of cutaneous secretory reactions, *J. gen. Psychol.*, 11, 1934, 445-448.

are: (1) a plastic chamber for enclosing a circular area of the palm and for providing a path by which the gas is passed over the palmar surface and then through the cell; (2) a source of dry nitrogen gas with equipment for the control of pressure and flow-rate; (3) a D.C.-power-supply to provide the electrolysis-current; and (4) a D.C.-power-amplifier and recorder.

Plastic chamber. Fig. 1 shows the cell and the plastic chamber mounted on *S*'s hand with the gas-supply and power-input cable connected. Fig. 2 shows the

FIG. 1. SWEAT-RECORDER MOUNTED ON *S*'S HAND

construction-details of the chamber. The main portion of the chamber is machined from a single piece of plastic rod. The lower portion of the rod is machined to provide gas flow on to the palmar surface by way of four holes evenly spaced at the edge of the skin-area enclosed by the chamber. The gas leaves the skin-surface by way of a hole in the center of the enclosed skin-area and passes up to the cell. A by-pass outlet is provided as a result of which all the gas leaving the skin does not go through the cell. The reason for this by-pass is described below.

The upper portion of the rod is machined to provide a chamber into which the body of the cell is placed to hold it in a constant, upright position. This portion of the chamber is machined to form a reasonably tight fit with the body of the cell, permitting the cell to be held in a fairly rigid way without the use of a clamping device. The path of the gas-flow from the skin is coupled to the input of the cell by means of a plug that fits up into a hole in the cell-body. This plug is made of nylon rod and machined separately and glued into the upper portion of the chamber. Its exact dimensions are not given as the plug must be machined to form a fairly tight seal with the hole in the cell to avoid leakage of the gas. The gas leaves the cell by way of another hole in the cell-body (on the same end

as the input-hole) and escapes by way of a side port in the upper portion of the chamber.

Electrical contact with the cell is made by way of a miniature phone-jack mounted on the side of the chamber. Each side of the jack is soldered to a silver wire that runs the width of the upper portion of the chamber. Spring-contact electrodes on the cell make contact with the silver wires when the cell is put in place in the upper portion of the chamber.

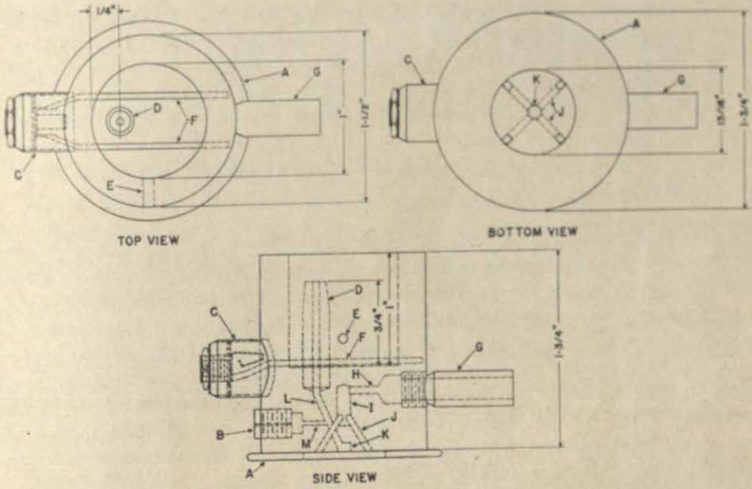


FIG. 2. CONSTRUCTION DETAILS

A, brass base-plate of chamber (1/16 in. thick); *B*, plug that screws into the side of the chamber and contains the by-pass outlet (0.028 in. in diameter); *C*, hollow plastic rod glued to side of chamber to house miniature phone-jack for making electrical contact with cell; *D*, plug that fits up into cell to allow a path for gas-flow from the skin-surface up into the cell; *E*, side-port in upper portion of chamber to allow gas to escape after passing through the cell (1/8 in. in diameter); *F*, silver wires (1/16 in.) to make electrical contact between the cell and the electrical input-jack; *G*, steel tube that is screwed into side of chamber to allow gas-input to chamber (5/16 in. in over-all diameter); *H*, hole from gas-input tube (1/16 in.); *I*, hole for gas-flow from gas-input to four holes leading to skin-surface (1/16 in); *K*, hole for gas to leave skin-surface (1/8 in.); *L*, hole for gas-flow up into cell (1/16 in.); *M*, by-pass hole leading to by-pass plug (1/32 in.).

The base-plate of the chamber is of brass. Brass was used instead of plastic as it appears to make a better seal with the skin and because it can be used as an electrode for skin-potential recording. It is attached to the base of the chamber by means of three screws (which, for simplicity, are not shown in the drawing). A terminal for connecting the base-plate to a lead-wire is mounted on the side of the chamber.

The top view in Fig 2 shows the dimensions of the top of the chamber, the approximate location of the plug that fits into the cell, and the location of the

silver wires. The side view shows the side-dimensions of the chamber and its internal construction. The bottom view shows the dimensions of the base-plate, the circular area enclosing the skin, the location of the four holes passing gas over the skin, and the location of the center-hole where the gas leaves the skin.

Source of dry nitrogen gas. Dry nitrogen gas is commercially available in drums of various size. A 240-cu. ft. drum has been found most convenient. The gas-flow control equipment consists of two units: a valve for pressure-reduction and control, and a flow-rate control-meter. On the input-side of the pressure-reduction valve is a gas-volume gage (0–300 cu. ft. range) that registers the gas-volume in the drum. The output of the pressure-reduction valve is coupled with a gas-pressure valve (0–50 lb./sq. in. range) to register the output pressure. The reduction-valve is adjustable to provide any output-pressure within its range. The flow-meter has a range of 25–700 cc./min. with a needle-valve control. A plastic tube ⅜ in. diameter connects the output of the flow-meter to the input-tube of the chamber.

Power-supply. The cell is designed to operate on 75–90-v. D.C. To protect the cell, a current-limiting resistor of 1 K is used. Initially, a battery supply was used with satisfactory results except that battery-replacement became troublesome. Later a power-supply was constructed with an output of 82-v. D.C. and a maximal current of 80 ma. In this circuit, a Sarkes Tarzin silicon diode (type M-150) serves as a half-wave rectifier with the output-voltage regulated by means of a Hoffman Zener diode (1N1373). The output-ground is isolated from the power-line ground by use of a 1–1 isolation transformer. With these components, the fluctuation in output-voltage is less than 1 v. over the entire range. All recordings reported here were made with this power-supply.

Recording equipment. The current through the cell is measured as a voltage-drop across a resistor in series with the cell. The resistor is coupled across the input to a D.C.-power-amplifier of an Offner Type-D *EEG* apparatus. A back-EMF is placed in the input-circuit to balance the input to the amplifier and to measure the voltage-drop across the resistor. Fluctuations in the voltage-drop across the resistor (being a function of fluctuations in current-flow through the cell) are then recorded by the ink-writer.

The size of the resistor and the range of the back-*EMF* depend on the sensitivity of the power-amplifier used. With the Offner amplifier, a 10-ohm resistor with a back-*EMF* range of 0–100 mv. (measurable in 1.0-mv. units), provide sufficient sensitivity for detecting sweating and range sufficient to encompass the variation in basal sweating levels. By Ohm's Law, the voltage-drop across the resistor is converted to amperage-units drawn by the cell. Thus, with the back-*EMF* measured in 1.0-mv. units, the amperage can be measured in 0.1-ma. units.

Gas-pressure and flow-rate. Gas-pressure, flow-rate, and temperature are important variables in the measurement of the moisture-content of the gas passing through the cell. A gas-pressure of 31 lb./sq. in. is high enough to provide a consistent gas-flow and low enough that the seal of

the chamber to the skin does not allow leakage. A flow-rate of 700 cc./min. is great enough for the detection of sweating responses but not great enough to overly dry the skin and thereby prevent the measurement of the basal sweating level. No attempt is made to control the gas-temperature. The temperature of the gas as it enters the chamber is 74 ±1° F.

With the gas-flow required for recording the response, a fairly high back-pressure, built up in the chamber, resulted which tended to produce leakage at the seal with the skin. To reduce the back-pressure, a by-pass outlet 0.028 in. in diameter was provided. This value was selected by trial and error. It reduces the pressure satisfactorily and at the same time enough gas is passed into the cell for reliable measurement of reaction.

Procedure. The operation of the apparatus to record palmar sweating is as follows: The chamber is strapped to the palm by means of two rubber straps that are hooked on either side of a ring clamp fastened to the upper portion of the chamber (Fig. 1). The hand is held palm up and against the body. The gas is then turned on for about 10 sec. and then turned off. If *S* reports feeling a slight pressure build-up and then its decrease when the gas is turned off, it is assumed that an adequate seal to the skin is made by the base-plate of the chamber. Then the gas is turned on again and set at 700 cc./min. With the back-*EMF* source adjusted to provide maximal back-voltage, the input to the amplifier is switched in across the series-resistor. With most *S*s, the voltage-drop across the resistor will still be higher than this back-*EMF* level, but in about 2 min. it will drop rather rapidly and assume a fairly steady level. The back-*EMF* required to balance at this level is noted and marked on the chart as the initial basal sweating level. This base-level represents a balance between the rate of moisture appearing on the skin-surface and the drying rate of the nitrogen gas. Discrete stimuli can be presented to *S* and sweating responses recorded as transient increases in the basal level. Voltage values are later converted to ma.-units.

This apparatus has been used thus far to record base level and responses of palmar sweating with 40 *S*s. Of these *S*s, only three failed to show consistent sweating responses in mental multiplication and tests of word-association. A record of sweating responses recorded from one *S* is shown in Fig. 3. The base level (in ma.-units) is marked on the record. The 1.5-ma. calibration mark shown on the right side of Fig. 3 can be used to measure the amplitude of the responses. These responses were elicited by problems in mental multiplication (audio recording shown on the lower tracing).

Latency of the apparatus. The latency of the sweat-recorder depends on two things: the latency of the cell and the time it takes the gas to reach the cell from the skin-surface. According to the manufacturer's specifications, the former is 0.7 sec. The latter can be computed from the volume of the path the gas follows between the skin-surface and the cell and the rate of gas-flow. The volume of the gas-path was computed to be 0.88 cc. With the rate of flow of 700 cc./min., it would take the gas 0.076 sec.

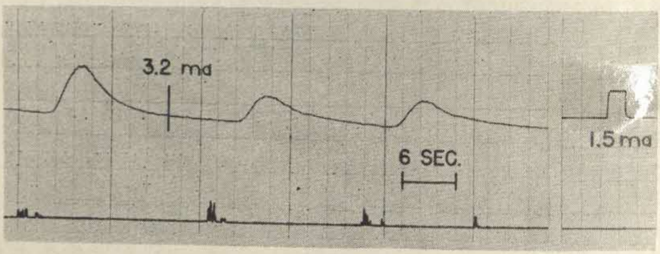


FIG. 3. RECORD OF PALMAR SWEATING

to traverse the gas-path from the skin to the cell. Thus the latency of the sweat-recorder is 0.776 sec. according to these calculations.

Response as a function of basal level. For the 37 *S*s, the relation between the basal level of sweating and magnitude of response was investigated. For each *S*, the amplitude of the first 10 responses was measured and the mean computed. The basal level 1 sec. before the onset of each response was measured and the mean computed. A plot of the data showed a positive, somewhat curvilinear relation between response-magnitude and basal level, the Pearsonian correlation being 0.71 and Eta being 0.84.

Modification of the apparatus. The apparatus as described appears to function quite well but improvement always is possible. One modification which has been made is to mount the cell on the wrist with only the chamber strapped to the palm. With this arrangement, a plastic tube is used to connect the chamber to the cell for the gas-flow path. The main advantage of this modification is that the chamber is less cumbersome to strap to the palm (it can be made quite a bit smaller) and it is not necessary for the *S* to keep his palm in an upright position. The disadvantage is that the response-latency of the apparatus is somewhat greater (being a function of the size and length of the connecting tube).

NOTES AND DISCUSSIONS

FRANCIS HUTCHESON AND THE THEORY OF MOTIVES

Whether all behavior is selfishly motivated was one of the most controversial psychological questions of earlier centuries. Adherents of the 'self-love' view included such men as Thomas Hobbes, Jeremy Bentham, and J. S. Mill. Even acts generally judged altruistic were considered selfish under this view. The 'self-love' orientation appears to underlie the concepts of some more recent psychologists. The earlier Freud (the pleasure-principle) and L. T. Troland (hedonism of the past) are out-and-out hedonists and, consequently, 'self-love' theorists. Both Thorndike and Hull should probably be thought of as more subtle hedonists. For example, Hull writes "Bentham's concept of *pain* is equated substantially to our own concept of *need*. . . . Bentham's concept of pleasure is found in those situations in which need or anxiety . . . is in the process of reduction."[1]

In view of the existence of such thinking in psychology, the arguments of Francis Hutcheson, a little known 18th-century philosopher, against the 'self-love' theory are important.[2] The kinds of distinctions which Hutcheson makes are involved in the statement of the problems of modern motivational theory. At the close of the paper I shall try to make explicit the connection of Hutcheson's remarks to contemporary theory.

The 'self-love' theory takes two forms: psychological hedonism and psychological egoism. Hedonism, which Hutcheson explicitly discusses, holds that all behavior is motivated by the actor's desire to attain pleasure or avoid pain. Egoism will be defined and discussed later. Hutcheson saw that the plausibility of hedonism derived from the ambiguity of 'pleasure' and 'pain.' Consequently, he distinguished three kinds of pleasure-pain and argued that none could be the sole motive of behavior.

First, present accompanying pain is the sensation which attends pressing desires.[3] Can the hedonist claim behavior is motivated by the desire to remove a present accompanying pain? Hutcheson's answer is 'no.' The pain which accompanies a desire for an object cannot be the motive for trying to get the object, since the pain was called into existence by the already existing desire for the object. We may add that his argument is

[1] C. L. Hull, *A Behavior System*, 1952, 341.
[2] Francis Hutcheson, *An Essay on the Nature and Conduct of the Passions and Affections: With Illustrations on the Moral Sense*, 1756.
[3] *Ibid.*, 15-16.

not refuted even if the accompanying pain becomes intense enough to cause a desire to relieve *it*. Such secondary desires could not arise without primary desires.

Secondly, gratification-pleasure is pleasure which attends the *satisfying* of any desire.[4] This pleasure should not be confused with the pleasure derived from the object obtained. Gratification-pleasure cannot be a motive of behavior. If it were, we would rate one desire as important as another, as gratification itself would be the real goal; but we do rate some desires higher than others. It is not mere gratification, therefore, which moves us. To say that gratification-pleasures are the objects of desire reverses the psychological process. Hutcheson could point out, as with accompanying pains, that gratification-pleasure cannot be the motive to try to get an object, since one must first have the desire for the object before the question of gratification comes up.

Thirdly, future pleasure and pain is best explained by comparing it to gratification-pleasure.[5] Consider the desire for food: gratification-pleasure is derived merely from gratifying the desire; future pleasure is that which results from eating the food, *i.e.* the pleasant taste. Hutcheson's first argument against this version of hedonism is that desire cannot be affected by volition. Suppose a person believes that if he desired to drink beer, such a desire would lead to future pleasure; he would, then, desire to desire to drink beer, but still might not desire to drink beer. The argument holds for benevolent affections—if someone thought he would be happy if he were benevolent, he would not thereby become benevolent. It is misleading, however, to construe this fact as an argument against hedonism. Hedonists do not argue that we *choose* to desire something because we judge it to be conducive to future pleasure; they argue that we are determined to desire what we think is the largest future pleasure or the least future pain. He felt the inadequacy of his argument and added, "the best Defenders of this part of the Scheme of Epicurus, acknowledge that Desires are not raised by Volition."[6]

Hutcheson's second argument is cogent. Hedonism is a theory of the feelings as well as of overt behavior. If the hedonist is correct, then, if an agent judges that some person or thing will produce his own largest pleasure, he will love that person or thing—or *vice versa*. Hutcheson points out that the consequent (that we always love or hate a person according to the amount of pleasure or pain involved for us) is false. We love and hate independently of future consequences, often to our dis-

[4] *Ibid.*, 16-17. [5] *Ibid.*, 17-21. [6] *Ibid.*, 20.

advantage. We may add that hedonism also greatly oversimplifies behavior, making it out to be much more rational (calculating) than it is. An argument similar to Hutcheson's about present accompanying pain can be made here: one must first have a desire for some object before one will receive pleasure from having that object.

There is, then, no basis for saying that 'altruistic' desires are necessarily really only disguised and complicated ways of seeking one's own pleasure.

Unfortunately, Hobbes, probably one of the most influential 'self-love' theorists, was not a psychological hedonist but a psychological egoist. Egoism is a 'self-love' theory which asserts that all behavior and affections are motivated by desires for what is in the agent's interest. Hedonism is a *species* of egoism. Hutcheson argues in terms of pleasure and pain, which refutes hedonism, but seems to leave egoism untouched. His arguments, however, can be adapted against egoism.

Consider the transformations that *hedonism* goes through in explaining a difficult case such as masochism. The hedonist must assert that the masochist's painful experiences are pleasant. This assertion has a certain plausibility, attributed by Sidgwick, the 19th-century English philosopher, to ambiguity of the words 'pleasure' and 'pain.'[7] (I shall discuss only 'pleasure' since 'pain' is similar.) 'Pleasure' may mean 'pleasant sensation' or 'that which pleases me.' Thus, the hedonist is not asserting (although he may not be clear about it) that the masochist finds the painful sensations to be pleasant sensations but that the painful sensations *please* the masochist, and this is true but vacuous. This vacuity can best be seen in an illustration. Jones goes to a psychotherapist, tells him the details of his history, and pays him large fees. When Jones has finished, he says, "Can you tell me why I persistently act in a way which alienates my friends?" The therapist says, "It gives you pleasure (pleases you) to act so—that is *why*." Jones should demand his money back, for the 'explanation' is theoretically useless (although it might have a therapeutic effect). It gives no information and no competent psychologist would give it. To say that painful sensations *please* the masochist and that is *why* he acts as he does says nothing.

The hedonist has two theories, not one, and the two are confounded. One asserts that behavior is motivated by desires for pleasant sensations, etc. Masochism, for example, disproves this theory. The other theory asserts that behavior is motivated by desires for what pleases, etc. The

[7] C. D. Broad, *Five Types of Ethical Theory*, 1951, 186-187.

latter is not a theory at all, but true by definition. Of course, one does what one pleases but the important question is, "What things please?" In argument the hedonist shifts from one theory to the other.

Egoism asserts that behavior is motivated by desires for what is in the agent's interest. A masochist would ordinarily be described as acting contrary to his own interest. This case invalidates a simple, straightforward egoism. To accommodate such cases, the egoist must argue that a masochist has two kinds of interest: Interest A, which is to act contrary to his own Interest B. Interest B must mean something like 'pleasant sensations,' 'happiness,' 'agreeable feelings,' *i.e.* the agent 'gets something for himself.' The egoist's argument, however, depends on *Interest A* and it has the same vacuity that *that which pleases* has and, consequently, egoism, if true, is identical with the vacuous interpretation of hedonism. To act from Interest A is merely to act from some motive or other, in the same way that *to do what pleases* is to act from some motive or other. Saying a person acts from an Interest A says no more than that the person has a motive. The important thing, however, is to know what kind of motive the person has. Hutcheson's arguments, therefore, are also effective against egoism.

It would have been amazing, if a pair of terms such as 'selfish' and 'altruistic,' which have such a long history of use, had turned out to be empty of content because *all* acts are selfish.

Commentators usually misleadingly refer to the dispute about motives as a matter of psychology, referring to "Hutcheson's psychology," etc. Psychology is an empirical science involving hypotheses, generalizations, measurement, observations, etc.—in short, substantive conclusions. But the 'theories' of Hobbes and Hutcheson do not meet these specifications. The facts they cite are known to everyone: affectionate feelings; sacrificial acts; admiration of the virtues of an enemy, etc. There is no dispute over these facts, as facts. They differ in their analyses of these common data. One concluded *all* acts are selfish and the other that *some* acts are altruistic; this is not, however, an empirical difference. The variant conclusions are due to the analyses that each gives to certain central concepts; namely, *interest, pleasure,* and *pain.* Hutcheson's contribution concerning motives is properly thought of as philosophy of science, *i.e.* as the analysis of the meanings of words, rather than as empirical science.

The clarification of the pleasure-pain concept is an instance of the kind of theoretical task which any complex scientific task involves. This linguistic task is non-empirical in the sense that what is involved is the analysis of the meaning(s) of central concepts—no substantive matters are involved. In fact, substantive conclusions cannot be safely drawn until

one is clear about the meanings of the concepts which are used to state the conclusions. If the analytic job is neglected, 'substantive conclusions' may turn out to be tautologies similar to "one always does what one pleases." Perhaps Freud's use of 'pleasure' involved such a vacuity; certainly John Stuart Mill's use of 'pleasure' in stating his psychological conclusions did. Thorndike's well-known struggles with circularity in trying to specify behavioristic criteria for satisfying and discomforting states is another example of the linguistic difficulty involved in stating conclusions about motives.

Washington State University GEORGE DICKIE

ATTENTION AND THE KINESTHETIC FIGURAL AFTER-EFFECT

In a recent paper on the role of attention in perception, Köhler and Adams concluded that attention may increase the articulation of an object and thus enhance satiation.[1] Their experiments also provide some support for the notion that the articulation of one set of functional characteristics in the psychophysical field presupposes the suppression of others. Further evidence on this point is presented here.

The *O*s were 24 students (16 men and 8 women) from graduate and undergraduate courses given by a variety of departments at the University of Buffalo. All were naïve as to the purpose of the experiment.

The apparatus used was similar to that of Klein and Krech, consisting of a test- (*T*-) object 1.5 in. wide by 6 in. long, an inspection- (*I*-) object 2.5 in. wide by 6 in. long, and a comparison-scale, 30 in. long, which was tapered from 0.5 in. at the narrow end to 4 in. at the wide end.[2] Each of the three objects was equipped with a 'rider' which fixed thumb and forefinger opposition.

The *O* was seated at a table and blindfolded. The *T*-object or *I*-object (as the case might be) always was presented to the right hand and the comparison-scale always to the left, since handedness does not appear to be of importance in this task.[3] Four practice-judgments of the *T*-object, two ascending and two descending, first were made. Next there were four *presatiation* judgments (again two of each kind) to provide an estimate of the *PSE*. Each judgment required 5-10 sec. The *I*-object then was presented to *O*'s right hand while his left hand rested on the table. He rubbed the sides of the *I*-object with his thumb and forefinger at an even rate (approximately 2-3 strokes per sec.) for 60 sec. Four more (*postsatiation*)

---

* This research was supported, in part, by the U.S. Army Medical Research and Development Command, Department of the Army, under Contract No. DA-49-007-MD-866. The work was done at the University of Buffalo.

[1] Wolfgang Köhler and P. A. Adams, Perception and attention, this JOURNAL, 71, 1958, 489-503.

[2] A diagram of the apparatus is presented in G. S. Klein and David Krech, Cortical conductivity in the brain injured, *J. Personal.*, 21, 1952, 126.

[3] Wolfgang Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesis, *Miscellanea Psychologica Albert Michotte*, 1947, 196-220.

judgments of the *T*-object followed this inspection-period. After a 5 min. rest-period, the entire procedure was repeated with the exception of the practice-judgments.

The two presentations permitted each *O* to serve as his own control in the experiment. Under the *control* condition, *O* was instructed merely to rub the *I*-object during the inspection period. During the *experimental* condition, a digit-span test was administered while *O* rubbed the *I*-object, *i.e.* *O* had to repeat series of digits which were read by E. For half the *O*s the *control* condition came first, while for the remaining half the *experimental* condition came first. This ordering of the two presentations also was followed for the subgroups of men and women in order to control for any sex-differences which might exist.

The measures of the kinesthetic figural after-effects were calculated by expressing the difference between the mean presatiation judgment and the mean postsatiation judgment as a percentage of the mean presatiation judgment. The mean after-effect scores were 8.92 for the control condition and 5.29 for the experimental condition, the difference between the two means being significant beyond the 1% level (t for correlated measures = 3.57). This result appears to support the hypothesis under scrutiny, since the conditions of the experiment were such as to rule out an explanation in terms of differences in stimulation.

As has already been noted, rate of movement in the inspection-period was held constant. Furthermore, the apparatus permits little variation in the pressure of the fingers, and pressure has been found to be of little or no influence. The important variable seems to be the position of the fingers, and this factor also was controlled. We should conclude, then, that when attention is distracted from an object, the articulation of the object and the resulting satiation are reduced.

Washington University
University of Buffalo

NORMAN L. CORAH
WALTER COHEN

A STATISTICAL TEST OF CYCLICAL TRENDS, WITH APPLICATION TO WEBER'S LAW

A search of the psychological journals revealed no study in which statistical procedures were employed for deciding whether to reject Weber's law. All decisions concerning the tenability of Weber's law were made on qualitative bases. This article describes the application of a statistical test that is appropriate for testing Weber's law.[1]

[1] The statistical test employed here and the associated probability table are presented in a separate publication, E. S. Edgington, Probability table for number of runs of signs of first differences in ordered series, *J. Amer. stat. Assn.*, 56, 1961, (No. 293), 156-159. The test is a modification of a previously published test, W. A. Wallis and G. H. Moore, *A Significance Test for Time-Series and Other Ordered Observations*, Technical Paper 1, National Bureau of Economic Research, 1941.

Since it is generally agreed that Weber's law does not hold for extreme intensities, only data for the middle range of stimulus-intensity were analyzed statistically in this study. Data on discrimination of visual brightness were used, viz. König's data for white light.[2] These data were particularly appropriate for the statistical demonstration since they have been used by some investigators for showing the tenability of Weber's law and by other investigators for showing the untenability of that law.

The test employed was chosen because it can utilize quantitatively the aspects of the data that other investigators have used qualitatively in

TABLE I

Visual-Brightness Discrimination for White Light

(Data in millilamberts)

Log intensity	−1.70	−1.40	−1.10	−0.70	−0.40	−0.10	0.30	0.60	
Weber-ratio	.0560	.0455	.0380	.0314	.0290	.0217	.0188	.0175	
Sign of difference		−	−	−	−	−	−	−	+
Log intensity	0.90	1.30	1.60	1.90	2.30	2.60	2.90	3.30	3.60
Weber-ratio	.0178	.0176	.0173	.0172	.0170	.0191	.0260	.0266	.0346
Sign of difference		−	−	−	−	+	+	+	+

rejecting Weber's law: the gradual drop and gradual rise of the Weber-ratio curve.[3] This test tests the hypothesis of random fluctuation of the numerical values in a series against the alternative hypothesis of systematic rises and falls in the numerical values in the series. It deals with the direction of a change in successive numerical values in a series and disregards the magnitude of a change.

The signs shown in Table I indicate an increase (+) or decrease (−) in the size of the Weber-ratios as the stimulus-intensity increases. The test involves counting the number of runs of signs of successive differences, where a run is defined as a group of one or more identical signs. When numerical values are randomly arranged in a series, one can expect considerable fluctuation and, consequently, several runs; when numerical values vary systematically in the series one usually expects few runs. The 17 Weber-ratios shown in Table I yield 4 runs: a run of 7 minuses, followed by a run of 1 plus, then a run of 4 minuses, and finally a run of 4 plusses. The probability-table for the test shows the probability of getting as few as 4 runs with 17 randomly arranged observations to be less than 0.0001. This result indicates a distinct tendency for the

[2] J. Blanchard, The brightness sensibility of the retina, *Phys. Rev.*, 11, 1918, 81-99.

[3] E. G. Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 136-139; A. H. Holway and C. C. Pratt, The Weber-ratio for intensity discrimination, *Psychol. Rev.*, 43, 1936, 322-340.

Weber-ratio to vary systematically with changes in stimulus-intensity; the decision is to reject the hypothesis that Weber's law applies over the middle range of stimulus-intensity. Statistical support now has been lent to a decision previously based on qualitative evidence alone.

Another type of application of this test can be made to examine the unreliability of observations repeated over a period of time (whether the time-interval between successive observations be a millisecond or a decade), when the observations are listed in temporal sequence. Such an application of the test could confirm the existence of a cyclical process of gradual growth and decline in visual acuity, intelligence, anxiety, reaction-time, or any other characteristic. Even measurements from very arbitrary measurement-scales can be utilized, provided the measurements can be arranged in rank-order of magnitude.

Kansas State Teachers College
Emporia, Kansas

EUGENE S. EDGINGTON

THIRTY-SECOND ANNUAL MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The Eastern Psychological Association met April 7 and 8, 1961, at the Bellevue Stratford and Sylvania Hotels in Philadelphia. There were 1950 persons registered at the meetings of whom 1110 were members of the Association, 285 were new members who joined the Association at the meeting, and 585 were guests. The present active membership of the Association totals 3360.

A committee headed by Norman Gekoski was responsible for local arrangements and the program was planned by a committee under the chairmanship of Raymond Katzell. The program consisted of 276 papers (presented in 54 sessions), 10 symposia, 5 special meetings, and 1 film.

Special meetings included in the program were: an organizational meeting of the Society for Client-Centered Counseling, a meeting of Psi Chi, a workshop of state association officers, the Princeton Luncheon, and a meeting of the Society for the Experimental Analysis of Behavior.

S. S. Stevens presented the annual presidential address, "The nomological quest." During the business meeting it was announced that the following new officers and directors had been elected: President, George A. Miller, and Directors, James Deese, James J. Gibson and E. J. Shoben. Marvin Iverson was elected by the membership to the post of Secretary for a three-year term. It was also announced that the Board of Directors had voted to establish the office of Historian and had appointed Gorham Lane to fill that post.

The 1962 meeting of the Association will be held at the Chalfonte-Haddon Hall in Atlantic City, April 27 and 28. The 1963 meeting will be in New York.

New York, N.Y. CARL H. RUSH

ERRATA

Two errors in the use of symbols occurred in J. W. Kling and Harold Schlosberg's article on "The uniqueness of patterns of skin-conductance," which appeared in the March, 1961, number of this JOURNAL (Vol. 74, pp. 74-79). In line 7 of the paragraph on *Methods and procedure* (p. 75) and in line 3 of the last paragraph on p. 77, the symbol μG (micromhos) should have been used in place of μΩ (microhms) of which μG is the reciprocal.

Roland Clark Davis: 1902-1961

The death of Roland Clark Davis on February 23, 1961 marks the end of major phases of the growth both of the Department of Psychology at Indiana University and of research on bioelectric phenomena in American psychology. Despite his fragile health, which had long been a matter of serious concern to his family and close associates, there was scarcely a day during a period of more than thirty years in which he did not contribute full measure toward bringing his institution and his science to their present stature.

Roland Davis' unusual record of identification with a single line of research is scarcely presaged by the pattern of his early life. He was born in Cambridge, Massachusetts, on December 20, 1902. His father, William Chalmers Davis, was a man of considerable intellectual reputation, an educator rather than a scientist. Evidently Roland's interests during school years leaned toward scholarship, but not especially toward science. After his graduation from the Easton, Pennsylvania, high school and two years of study at Lafayette College, he went on to Harvard to receive his A.B. in 1924 with a major in English. It would be interesting to know just what factors prompted his apparently abrupt shift of interest. After a single year spent teaching English in the high school in Greenville, Ohio, he enrolled for graduate work in psychology at Columbia in the summer of 1925, and, though somewhat interrupted by a period spent earning money as a school psychologist in Petersburg, Virginia, he continued to the Ph.D. at Columbia in 1930.

It is not hard to identify two critical events during the period of

Roland's life which he entered as a student of the humanities and from which he emerged as one of this country's first full-time psychophysiologists. The first of these was meeting and marrying a lovely young school teacher, Frances Meacham, during his otherwise unproductive stay in Petersburg. Beyond the companionship and support which plainly meant so much to him throughout the rest of his life, his wife contributed immeasurably to the warmth and gaiety of the Davis home in Bloomington, where many generations of incoming faculty and graduate students were made to feel that they had joined a community rather than merely an institution.

The second turning point of the transitional period was, of course, Davis' coming under the influence of Woodworth and Poffenberger at Columbia. It would be difficult now to apportion the contributions of these two eminent teachers in molding his approach to psychology. Certainly Poffenberger must have had the more important hand in completing Davis' conversion to the way of the laboratory scientist. In any event, the conversion clearly was complete by the time he wrote his doctoral dissertation on "Factors affecting the galvanic reflex."

After completion of graduate study at Columbia, he spent two years as a research associate at the University of Virginia, then went to Indiana University in the summer of 1931 as Acting Associate Professor of Experimental Psychology. This appointment was made permanent the next year and was followed by promotion to a full professorship in 1938.

Friends, colleagues, and students have arranged for a portrait of Roland Davis to overlook the entrance hall of the new building which will shortly house the psychological laboratories at Indiana University. It will be a fitting reminder of the many years during which Davis' familiar, slightly stooped figure was to be seen in the halls of the present building seven days a week, twelve months a year, showing students by example rather than exhortation that scientific research is a manner of living rather than a job to be conducted during bankers' hours.

Appearing in sharp relief against his early literary background, Davis' enduring concern with problems of instrumentation in psychology emerges full-blown in the reports of his earliest researches. His first published article, appearing while he was still a graduate student, reported a new vacuum-tube technique for stabilizing current during measurement of the galvanic reflex. His doctoral dissertation, with the thoroughness and meticulous attention to detail that were to become almost a trademark, begins with a critical survey of the techniques used in some four hundred

studies of the galvanic reflex that had been reported during the preceding thirty years, then proceeds to spell out the requirements for proper measurement on the basis of physico-chemical theory. This essay was followed by a long series of contributions to instrumental methods for the measurement and recording of the galvanic reflex, muscular tension, muscle action potentials, and circulatory changes, and by a number of general reviews of this area, including his "Methods of recording action" in Andrews' *Methods of Psychology*. His work in psychophysiological measurement and recording was unusual in that he viewed its problems, not as matters to be handled simply by clever gadgetry, but as a serious branch of research, requiring the founding of techniques in physical and physiological theory, and with special attention to the effects of measurement-operations on the process being measured.

Psychologists familiar with Davis' professional life-long absorption in psychophysiology exhibit surprise at seeing listed among his publications the book-length *Ability in Social and Racial Classes*. The apparent paradox is soon dispelled, however, when one notes the subtitle, *Some physiological correlates*. This volume, the product of his research associateship at the University of Virginia, reports his study on the relationships between measured intelligence and properties of relatively simple reflexes, and the correlation of these with social and racial variables. He did indeed find evidence for patterns of correlations suggesting a systematic relationship between metabolic level and measured intelligence. Of far more import for his later work, however, was a byproduct of the study. As one of the biological measures, he had chosen the latency of the Achilles reflex, with the thought that this would provide the most direct available index of speed of nerve-conduction. Seeking a means to eliminate time of effector-action from the latency-measure, he concluded that the way was to record the action-potential given off by the muscle just prior to its actual contraction. After thorough analysis of the methods used previously by Travis, Hunter, and others, he arrived at the technique of recording the muscle action-potential by means of a cathode ray oscilloscope, a method which came to be basic to much of his later work.

To all but a handful of fellow professionals, the history of Davis' thirty years of research on the role of muscle action-potentials in psychology is hard to comprehend. He had developed an apparently satisfactory method of measurement while searching for the biological basis of intelligence in the early 1930's. Then for nearly a quarter of a century, a steady output of technical articles and memoranda reported a long series of researches,

meticulously conducted but extremely limited in scope, dealing for the most part with recording of responses to simple stimuli under the simplest of laboratory conditions. Suddenly, during the 1950s, Davis authored a series of articles of wide theoretical interest in which he applied his techniques to the problem of recording and measuring the neurophysiological variables hypothesized in theories of learning and motivation. One of these applications dealt with the stimulus-trace, basic to Hull's treatment of conditioning, another with detection of the covert correction-response postulated in behavioral theories of multiple-choice learning, a third with the conditioning of passive movement. A series of provocative studies concerned with the electrical correlates of Cannon-Carlson waves and hunger-pangs seems likely to lead to radical revision of classical views of the neurophysiological basis of hunger. Picked up by other investigators, Davis' methods are currently being used in attempts to secure independent measurement of the anxiety-drive which plays a critical role in some current interpretations of human learning.

Why, the student wonders, upon seeing the numerous and fertile applications of Davis' electrical recording techniques, did all this not come about much earlier? The answer, of course, is that, as is so often the case in science, major problems of interpretation had to be solved before muscle action-potential recording could be brought to bear upon problems of general interest. The complex and variable record obtained when one leads electrodes from the surface of the skin to an oscilloscope bears no simple relation to what is going on beneath the surface. Trying to trace the activity of designated effector-units by means of these recordings is much like trying to follow an athletic contest by listening to the noise of the crowd from outside the stadium. Three decades was all too short a time in which to bring this line of research to fruition.

In one of his rare excursions into theory, Davis presented his conception of "a physical psychology" in the *Psychological Review* in 1953. His view was that psychology should be *part of* physics, not merely *like* physics. The postulating of entities to explain behavior he considered to be not sophisticated methodology but simply shoddy workmanship. Hypotheses concerning physiological events are all right for a start, but they are of no real value until ways are found to make these events observable and measurable. Whether or not behavior theorists come to accept the proposed criterion of "observation of the thing hypothesized instead of merely the observation of the behavioral consequences," there is no doubt but that they will benefit from his having so eloquently expressed it.

Indiana University

W. K. Estes

Carl Iver Hovland: 1912-1961

Carl Iver Hovland was born in Chicago on June 12, 1912, and died in New Haven on April 16, 1961. As a youngster in Chicago, he attended the Lloyd School and then completed high school at the Luther Institute. He entered Northwestern University at the age of 16, receiving his B.A. in 1932, and an M.A. the following year. He then transferred to Yale, where he obtained the Ph.D. in 1936. Except for a three-year research stint in Washington during World War II, Hovland remained associated with Yale the rest of his life, rising rapidly through the academic ranks to a Sterling Professorship at the age of 36.

During his high school and college days, Carl Hovland was deeply interested in music, and, until psychology caught his attention permanently, he looked forward to a musical career. He was an excellent pianist and organist, but once having shifted into high gear in psychological research, his playing became mainly a hobby for family enjoyment. Music did bring him his wife, however; in 1938 he married Gertrude Raddatz, a fellow piano student in Chicago. Their home was constantly filled with music, usually from one of the magnificent high-fidelity sets that Hovland was forever constructing—he was as knowledgeable in electronics as he was in music.

Psychological research, was, however, his chief intellectual joy. There was something about the application of experimental methods to behavioral phenomena that fascinated him. His investigations covered a wide range of problems, especially during his early career. By the time he secured his doctorate, he had published a dozen research papers and had collected the data for half as many more. Four of these 18 papers were in the *American Journal of Physiology,* two in the *Yale Journal of Biology and Medicine;* the others, in psychological journals, included a major review of the literature on apparent movement, a study of test-reliability, and his four now classical papers on conditioned generalization. These latter were from his doctoral dissertation.

His catholicity of interests continued during the next six years. He was one of the 'young Turks' brought into the Yale Institute of Human Relations by Mark May in 1936, and his research fell into four areas. Already stimulated by Clark Hull's attempts to construct a rigorous learning theory, Hovland immediately joined in the effort, and for several years produced a flood of precisely designed studies of rote learning. Simultaneously he was participating in the group that eventually published *Frustration and Aggression,* of which he was one of the collaborating authors, and in a smaller group working on conflict. In spite of his

extraordinary productivity in research during this prewar period, Hovland found time to carry on an active program as a consultant to two or three industrial firms, and to support a normal teaching load at Yale.

With his broad range of experience in research, it was not surprising that in 1941 he was one of the first psychologists chosen to serve in the then new Division of Information and Education in the Office of the Chief of Staff of the War Department. After the war, Hovland, Lumsdaine, and Sheffield published a report of the major research findings, from one part of this unique enterprise, under the title *Experiments on Mass Communication* (Princeton University Press, 1949). The war setting had provided an opportunity for some very large well-controlled experiments in the area of opinion-control, and the results were so rewarding that Hovland settled thereafter on the general field of communication as his main research interest.

Returning to New Haven in 1945, Hovland served for three years as Chairman of the Department of Psychology. The expansion of his program of research in communication finally forced him to make a choice between it and departmental administration and not surprisingly he chose the research. Although he remained interested in problems of verbal learning, and wrote a chapter for Stevens' *Handbook* as late as 1950, the main contributions of his last fifteen years were in concept-formation and communication. In the latter field, his interest lay in the conditions under which opinions or attitudes can be modified. His search took him from the personality of the communicator to that of the communicatee, from the order of presentation of ideas to the content of the ideas themselves. Perhaps one of the most important of his contributions was his analysis of the differences between the experimental and the survey studies of attitudinal changes, given in a paper before the American Psychological Association in response to the Association's Distinguished Scientific Contribution Award to him in 1957.

The other major interest of his last years was concept-formation, a field he approached from the standpoint of computer simulation. In 1952 he published a demonstration that the problems of concept-learning can be solved by an hypothetical decoding machine, and also provided both a notational system and an analysis of the problem of concept-learning that were widely adopted by other researchers in this field. This analysis made use of a newly developed mathematical theory of communication. A series of experimental reports by Hovland and his students soon followed, and, when the mid-1950's brought a major breakthrough in the design and programming of digital computers, Hovland quickly perceived the poten-

tial application of informational processing to psychology. He collaborated in the development of a computer program that would serve as a model of human performance, and in 1960 reported the first results of testing such a model.

Carl Hovland was a big man, soft in speech, gentle in manner, as incredibly quick and deft in physical movement as in intellection. In his earlier years he was quite shy, but the social rigors of the life to which his extraordinary talents inevitably exposed him helped to develop the quiet ease of manner that characterized his middle age. He was unfailingly cheerful, even in the last tragic year of his life, and continued to work with his students and colleagues until his brief final illness. It was this capacity for being always helpful, always objective, that placed him in constant demand as a consultant, not only to students, but to all the leading Foundations, to half a dozen major Government agencies, and to the behavioral research arms of several great corporations. He was repeatedly honored by his colleagues in one way or another—as an APA representative to the Social Science Research Council, as a member of that Association's Board of Directors, by election to the American Academy of Arts and Sciences, to the American Philosophical Society, and to the National Academy of Sciences. The honor he appreciated most deeply, perhaps, was the award of the Warren Medal by the Society of Experimental Psychologists, word of which reached him only a month before his death.

Stanford University ROBERT R. SEARS

Carl Gustav Jung: 1875-1961

On June 6, 1961, Carl Gustav Jung died at the age of eighty-five years in his villa in Kuesnacht, a suburb of Zurich, Switzerland. Born the son of a Swiss clergyman in Kesswil, on Lake Constance, on July 26, 1875, he gave early evidence of his later brilliance and wide-ranging interests. At the age of six years, he read Latin textbooks fluently, beginning the development of the linguistic skills which were to be so important in his later studies of alchemy and religion. From the beginning, however, he was known as much for his tremendous physical strength as for his mental prowess. In his 80s, his sailing and gardening continued, as did his almost unparalleled scholarly productivity. Truly, as Henry Murray observed at his first meeting with Jung, here was a "full-blooded, spherical, and Goethian intelligence."

As one of the founders of modern psychiatry, and the father of that branch of it called analytical psychology, Jung is revered, but, like much of modern art, he is more easily appreciated than understood. He who

rejected rigid categories in psychoanalytic theory and psychiatric practice is himself difficult to categorize after death. Still harder is it to assess what continuing influence he will have on an American psychology whose pet idols he so vigorously attacked.

For here was a psychiatrist who stressed experimental method, developing his Word-Association Test from earlier work by Wundt, but one who wrote for his M.D. thesis a book entitled *About Psychology and Pathology of So-Called Occult Phenomena.* Here was the son of a Protestant clergyman who became one of the world's foremost experts on Eastern religions, and who felt that Protestantism had much to learn from Rome. Here was the heir-apparent of Freud who became one of his most vigorous critics, with five years of intense intellectual interchange followed by twenty-six years with no contact at all. Here was one rejected by many of his colleagues as mystical or "too philosophic" who considered himself first and foremost an empiricist, condemning others for lack of proper attention to the full richness and complex interaction of "psychic facts."

Carl Gustav Jung remains, then, a widely honored and respected paradox. To his early honors as Bleuler's assistant and collaborator at the Burghölzli Mental Hospital in Zurich, and as first president of the International Psychoanalytic Society at the age of thirty-six years, were added many others, including honorary degrees from Harvard and Oxford Universities. His Word-Association Test, and the Sentence-Completion Tests which were developed from it, remain useful diagnostic tools. His theory of psychological types, which arose from his attempt to understand the differences among himself, Freud, and Adler, has passed from psychological testing into common usage, and even into school report forms which include the magic words *extravert* and *introvert.*

In a more general theoretical sense, his theory of emotional complexes—which the Word-Association Test sought to 'tap'—has had wide influence. In the studies of dissociation of personality by Janet, Prince, and others came empirical verification of it. Other unique contributions of Jung, in studies of mythology, Western and Eastern religions, and the process of individuation or self-actualization, have made great contribution. Nevertheless, comparatively little of Jung's insights and findings has been assimilated into American psychology and psychiatry. Freud's position has been too dominant for his major critic to receive much of a hearing.

In recent years, however, a new climate of potential receptivity to Jung's contributions has begun to emerge in America. This change is related to the current concern—or fad, if you will—for the phenomenal field and the growth of identity within it. As one who has regarded himself from the beginning as a phenomenologist, Jung has much to contribute at this

point to those who have sufficient patience and diligence to work carefully through his extensive writings—for, in study of the individual's developing selfhood, Jung takes more factors into account than do most contemporary theorists. He has, moreover, related these variables to one another in terms of a continuous, dynamic, consistent yet 'open-ended' process. He is concerned for future intention as well as the determinism of the past, for interpersonal "dialectical discussion" as well as internal structure and function. Most radically, he challenges causality as the universal principle of explanation and emphasizes "synchronicity"—"a peculiar interdependence of objective events among themselves as well as with the subjective (psychic) states of the observer or observers."

In this principle of synchronicity, and perhaps even more in his central concept of the archetypes of the collective unconscious, Jung still does not 'make sense' to most American psychologists. He is disturbing, and therefore valuable, precisely because he does force us to reëxamine many of our most cherished presuppositions. In his analytic practice and scholarly studies of symbolic productions, he has brought forth data which do not fit into the currently available categories of psychological theory and psychiatric practice. Thus, as Einstein challenged existing physical theories by using empirical data which did not fit into them, so Jung has challenged a personality theory which, in America at least, is based mainly on Freud's creative hypotheses. Jung—and in America, Allport, Rogers, Maslow, and others—has forced us to move beyond causal, impulse-driven, closed-system personology, as Freud blasted psychology out of static structuralism and associationism.

Perhaps Jung's most lasting contribution will be his criticism of psychological myopia and his furnishing data for future analysis and interpretation, rather than as a new system. That would please him, for he never sought disciples, or claimed to establish a 'school of thought' or systematic theory. His concern throughout his professional practice was that each individual experience and think for himself—indeed, *discover himself* in his complex wholeness, rather than identifying one portion with the whole. This, I am sure, would be his hope for personality theory as well.

Boston University WILLIAM DOUGLAS

Carl Murchison: 1887-1961

Carl Murchison died at the New England Baptist Hospital on May 20, 1961, after a relatively short illness. He was born in Hickory, North Carolina, on December 3, 1887.

He obtained the A.B. degree from Wake Forest College in 1909. His

career in psychology began when, as Johnstone Scholar, he received his doctorate from Johns Hopkins University in 1923. Almost immediately he became a notable figure in the profession by accepting the chairmanship of the department at Clark University as the successor of G. Stanley Hall.

Murchison's personal interest soon turned to publication. In 1924 he took over the editorship and management of the *Pedigogical Seminary*, adding to the title *and Journal of Genetic Psychology*. In that same year, 1924, he published his own book on *American White Criminal Intelligence*. The following journals he personally planned, founded, financed, edited, and managed under the name of The Clark University Press: *Genetic Psychology Monographs*, 1925; *Journal of General Psychology*, 1927; *Journal of Social Psychology*, 1929; and *Journal of Psychology*, 1935. In 1936 he left Clark University, set up The Journal Press in Provincetown, Massachusetts, and continued to edit and publish all of these journals up to the time of his illness and death. How carefully he planned them to fit the needs of the profession is attested by the fact that all of them survive.

Murchison also planned, edited, managed, and published the *Psychologies of 1925*; a book on *Criminal Intelligence* in 1926; *The Case for and against Psychical Belief*, 1927; *The Psychological Register* in 1928; *The Foundations of Experimental Psychology*, 1929; *Psychologies of 1930*; *A History of Psychology in Autobiography* in three volumes in 1930, 1931 and 1936; *A Handbook of Child Psychology*, 1931, second edition, 1933; *A Handbook of General Experimental Psychology*, 1934; and *A Handbook of Social Psychology*, 1935. Again his careful and penetrating analysis of the needs of the profession is shown in the wide acceptance and continued use of these works.

Murchison received honorary degrees from Wake Forest College and from the University of Athens. He was a member of Phi Beta Kappa. He collected some 250 paintings by Provincetown artists and a very large number of scientific treatises on psychology including, with very few exceptions, a copy of every psychological journal ever published both here and abroad. All of these things were lost in a fire at his home in Provincetown in 1956.

Florida State University

JOHN PAUL NAFE

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Psychology: A Study of a Science. Study 1. *Conceptual and Systematic.* Volume 3. *Formulations of the Person and the Social Context.* Edited by SIGMUND KOCH. New York, McGraw-Hill Book Company, Inc., 1959. Pp. x, 837. $12.50.

The present volume is the third in a projected series of seven under the general title, *Psychology: A Study of a Science.* This project was initiated by the Policy and Planning Board of the American Psychological Association. The contributors to the first three volumes (Study 1) were assigned the task of assessing current theoretical formulations of recent importance to psychology. The assignment was couched in broad terms—"theoretical formulations" being defined as "any set of sentences formulated as a tool for ordering knowledge with respect to some specified domain of events, or furthering the discovery of such knowledge." The contributors were furnished with an outline of suggested discussion topics—an outline which was carefully followed by some and ignored by others.

Within this framework the contributors were free to develop their respective fields of interest and research. Stated somewhat colloquially, it was as if the editor had said to a prospective contributor, "You have been working in ——— field for several years and have made substantial contributions to it. Now we would like you to summarize your work, explain its methodology, and expound its critical significance to psychology." Such an invitation is not too frequently met with. It is an invitation to brilliancy and creativity, and also, it must be noted, to prolixity which not all the authors unfortunately were able to resist. In the present volume the result is a series of independent essays, each self-contained, in which the author expounds the methodology and theory of his particular specialty.

Under the somewhat awkward title "Formulations of the Person and the Social Context" eleven specialists have written essays which occupy 711 pages, and the editor, Dr. Koch, has added a 53-page *Epilogue* which summarizes the major trends in all three volumes of Study 1.

In considering Volume 3 as a whole certain general impressions emerge. These may be expressed in two questions: What general purpose does the volume serve? To what extent is it representative of the area it is presumed to cover?

This reviewer has been unable to find a satisfactory answer to the first question. We have in effect eleven minor monographs in eleven widely separated fields of specialization. The editor specifically disclaims any intent to produce a grand integration or a *Summa Psychologica.* Nevertheless, the reader who examines the volume as a whole is bound to inquire for what audience are these expositions intended. In spite of the editor's expressed hope that the "self determining citizen" may read them to form his "own perceptions of the place of psychology in the pattern of modern knowledge," they are not addressed to the layman. They are, in fact, addressed to the specialist—specialists in eleven different fields, to be exact. But would not these specialists have found the same material in the numerous

professional journals and books to which these distinguished and highly articulate authors are frequent contributors? There are, of course, the graduate students. This volume and the others in the series, will be required reading, and doubtless will save the students some library work. Perhaps it will be the basis for stimulating seminar discussions.

As regards the second question, a not wholly satisfactory answer is suggested in Dr. Koch's *Introduction*. "Eleven topics" he notes, "is a stingy allotment relative to the range and density of effort in studies of the person and the social setting" (p. 4). It is indeed. No claim is made for representativeness. "We do," says Dr. Koch, "claim sufficient diversity to extend markedly the range of formulations which in recent years have been given sustained analysis." It is inescapable, however, that a work bearing the comprehensive title "Psychology: A Study of a Science" will be in part judged in terms of the extent a given field is represented. Probably no two psychologists concerned with psychology in the social context would agree as to the topics to be covered. For this particular reviewer there are some inexplicable omissions. The theoretic formulations and research methodology in the field of ethnic tensions would have been relevant to the announced purpose of the project. Not only because this field has highly important applications, but because it contributes significantly to our knowledge of man in society (the social context). Missing also is any reference to the theory and methodology of public opinion research. In these and other fields surely there is no paucity of formulations which "in recent years have been given sustained analysis."

The fields covered may be divided roughly into two groups: those concerned with "the person," and those concerned with the social context (traditionally called social psychology). In the first are Murray's "Preparation for a Scaffold of a Comprehensive System," Rapaport's examination of psychoanalysis, Rogers' presentation of the theory of client-centered therapy, and Cattell's exposition of factor analysis as applied to the study of personality. In the second group are Asch's perceptive on social psychology, Newcomb's study of attitudinal patterns, Katz and Stotland's study of attitude-structure, Lazarsfeld's presentation of latent structure analysis, Thelen's theory of group-behavior, and Parsons' theory of action.

There remains the essay on psychogenetics by Kallman which does not fit readily into either group. It is a competent and important statement of the theory and methodology of the study of twins. In view of the "stingy allotment" of space to the various fields, and in spite of the author's belief regarding the fundamental importance of his subject for psychology, it is difficult to understand the inclusion of this essay in this particular volume.

It should not be forgotten that by editorial intent each of these essays stands alone,[1] and it would be presumptuous for any one reviewer to attempt a critique of contributions in such widely divergent and highly specialized fields. In the present volume Dr. Koch discusses the major trends in all three volumes of Study 1. The reader who undertakes the not inconsiderable task of reading the present volume in its entirety would be advised to read Dr. Koch's *Epilogue* first.

As a minor supplement to Dr. Koch's brilliant synthesis, and a small assist to

[1] In fact, one of them apparently has already appeared as a separate monograph. *Contemporary Psychology* (Vol. VI, No. 4, 1961) carrys a review of David Rapaport's *Structure of Psychoanalytic Theory: A Systematizing Attempt,* reviewed by Thomas S. Szasz.

the "self determining citizen" who may attempt to read the volume, a brief characterization of the eleven essays seems to be in order.

In a "Preparation for a Scaffold of a Comprehensive System" Henry A. Murray presents what he terms an intellectual biography. He confesses his inability to follow the outline provided by the Editor, and his reasons are illuminating. He thinks it admirable "for exclusively experimental specialists, observers of closely restricted animal activities, peripheralists, positivists, but scarcely possible for naturalists, generalists, and centralists who study gradual transformations of the dispositions, beliefs, and modes of action of human beings as they manifest themselves in different social settings." (p. 8) In a word it is premature and generally inapplicable to the task of the book as he sees it. In his intellectual biography he traces the influence on the "scaffold" of his studies in medicine, of psychoanalysis, and of such men as Lewin, McDougal, Darwin, Kluckholn, Parsons, and Whitehead. Six conceptions are listed which have contributed to the scaffold: cathexis (from Freud), the dyadic relationship, thema, thematic dispositions, serials (time binding), and ordination (construction of plans of action).

David Rapaport in "The Structure of Psychoanalytic Theory: A Systematizing Attempt," does the best job that probably could be done with a set of concepts which are essentially unsystematic. He traces the influence on Freud of Brentano, literature, especially Goethe, and the general atmosphere of Victorian Vienna. He notes Freud's sensitivity to language-usage and his "knack for metaphor." He finds Freud's theory based on four models: the reflex arc (topographic model), entropy (economic model), the Darwinian (genetic) model, and the Jacksonian (integration hierarchy) model. He believes psychoanalytic theory possesses quantitative considerations but doubts if these are readily translated into actual measurements. He insists that the genetic formulations of Freud do not conflict with Lewin's ahistorical conceptions. He notes that although Freudian theory grew by "spurts," it possesses a structural unity which has been obscured by its numerous revisors.

Carl R. Rogers in "A Theory of Therapy, Personality, and Interpersonal Relationships, as Developed in the Client Centered Framework" also presents the biographical background of the development of his conceptions. He stresses the fundamental role of the subjective in his thinking, and voices the conviction that science does not necessarily begin in the laboratory or at the calculating machine, and the most helpful models are not derived from theoretical physics. He defines some forty key terms or phrases in his theory. These include the tendency towards self-actualization, experience, symbolization, self experience, incongruence between self and experience, vulnerability, and contact. The process of therapy is discussed in detail. Out of this process there is developed a theory of personality and the dynamics of behavior.

In "Personality Theory Growing from Multivariate Quantitative Research" Raymond B. Cattell is concerned with a theory based on quantitative and objective methods. He reviewed three methods in personality research: the clinical method, the classical univariate experiment, and the multivariate analytic experiment. Theory derived from the clinical approach based on diagnosis and therapy, he believes to be "in a bad way." The analytic multivariate method provides a swifter and surer approach to the significant variables for controlled experimentation. In developing an experimentally based personality theory the unimaginative classical univariate procedure has not proved useful. In psychology, as distinct from physics, it is neces-

sary to isolate organically unitary behavior structures before the univariate experiment can be fruitfully applied.

Franz J. Kallman's "Psychogenetic Studies of Twins" proceeds from the premise that twin-study is the only quasi-experimental procedure for the study of human genetics, and provides data fundamentally important for the study of human behavior. Procedural limitations and advantages of the method are discussed.

In a "Perspective on Social Psychology" Solomon Asch points out that social psychology still works largely with borrowed conceptions. General psychology historically has been restricted in its subject matter since it feared the complexities involved in relations between person, and persons and groups. Further bases for this restriction was the belief that fundamental psychological principles could be extrapolated from the findings of general psychology. The "curious doctrine that man is directly descended from the white rat" is rejected. Social psychology must accept the challenge that its phenomena are not amenable to the yard sticks of physics.

Theodore M. Newcomb in "Individual Systems of Orientation" examines the consequences of the postulate that there are lawful interdependencies among certain classes of beliefs and attitudes held by the same individual. The immediate concern of the essay is the systematic formulation of the processes by which human beings develop attitudes towards other persons and towards objects of joint relevance. These individual systems of orientation are developed primarily through communicative processes which are examined in some detail. Research in this field must utilize natural situations ordinarily not found in short-term laboratory experiments.

Daniel Katz and Ezra Stotland in "A Preliminary Statement of a Theory of Attitudes" discuss the affective, cognitive, and behavioral components of attitudes and their relation to value systems of individuals. The field-study approach to social behavior is emphasized rather than the narrower laboratory method. Formal mathematical models are regarded as making a very limited contribution to the progress of social psychology. In contra-distinction to physics, the barrier to advance in social psychology is the gap between genotypic constructs and phenotype measures. This is regarded as the basic difference between the social and natural sciences.

Paul F. Lazarsfeld in "Latent Structure Analysis" is concerned with the "more permanent reality" behind the "symptoms" which frequently are all that concerns the social scientist. If we read "phenotypic" for "symptoms" this approach would presumably narrow the gap that concerns Katz and Stotland. The theory and methodology of latent structure analysis are discussed in technical detail, including the presentation of the nine steps necessary in latent structure investigation.

In the "Work-emotionality Theory of the Group as Organism" Herbert A. Thelin reviews the historical background and theory of small group research. The type of face-to-face group studied is composed of twelve to twenty-five adults as found in the various human relations training "laboratories." In the study of such groups it is assumed that units of interaction exist, that personality factors reflect individual stresses, that group life is adaptive, that group factors exist by virtue of inter-active networks among individuals, and that the problems for investigation are found in the interaction between "personality" and the "group."

Talcott Parsons in "An Approach to Psychological Theory in Terms of the Theory of Action" breaks down his general theory of action into four "reference points": the organism, the personality or psychological system, the social system,

and the cultural system. These are sub-systems of action, and are not independent and not reducible, but are regarded as an order of levels. These sub-systems are discussed in detail. The author's personal and professional history is presented as a background for his theory. Gordon Allport, Murray, and W. I. Thomas are among those mentioned as influential in the author's thinking. He notes that he was troubled by the reductionist trends in much of behaviorist psychology, especially that of Watson and Hull. It is worth noting that Parsons, of the eleven contributors, is the only one who gives credit to the writings of George Mead. This is odd considering the basic character of Mead's hypotheses for several of the topics treated in this volume.

These thumbnail sketches make no claim to comprehensiveness, and they no doubt will seem insultingly inadequate to the respective authors. Although anything like a summary of content so diversified is impossible certain trends do emerge. Some of these are pointed out in Dr. Koch's brilliant *Epilogue*. Three impress this reviewer.

1. The methodology and orientation expressed by the S–R type of theorizing is viewed as inadequate for social psychology. Extrapolation from what Gordon Allport has called the "antics of captive and desperate rats"[2] is not regarded as a sound basis for the understanding of man in the social context.

2. There is concern, even nervousness, about the use of the intervening variable paradigm, although it is by no means abandoned. The need to close the gap between phenotypic observation and genotypic construct is expressed, although just how is not made clear.

3. Finally, with possibly two exceptions, mathematical models in social psychology are viewed with skepticism, although quantification is not to be abandoned. A closely related attitude expresses doubt regarding the adequacy for social psychology of the classical experimental laboratory procedures modeled on the methods of the exact sciences. Field methods and natural situations for the observation of social behavior are regarded as essential to social psychology.

In spite of occasional ponderosity of phrasing and the apparently almost compulsive need of each author to develop his own esoteric jargon, these essays, considered individually, are a remarkable achievement. The authors have met a challenge with competence, and in some cases with brilliance. Some are worthy of special mention. Because of its modesty, frankness, and general readability the essay by Murray is exceptional. The essays by Asch, Katz and Stotland, and Newcomb seem especially noteworthy, not only because of their lucidity and competence, but because to this reviewer at least, they seemed closer to the core of social psychology.

Considering the book as a whole, I am inclined to apply to this volume the estimate expressed by Ryan in his review of Volume 1,[3] "it is valuable for its excellent parts, but this is at least one instance in which the whole is not more than the sum of its parts."

FRANKLIN FEARING

University of California,
Los Angeles

[2] Gordon Allport, *Becoming,* 1955, 18.
[3] This JOURNAL, 73, 1960, 492.

The American Voter. By ANGUS CAMPBELL, PHILIP E. CONVERSE, WARREN E. MILLER, and DONALD E. STOKES. New York, John Wiley and Sons, 1960. Pp. viii, 573. $8.50 (trade edition), $6.50 (text edition).

This book is an interdisciplinary classic. Its merit as a contribution to political science has already been amply recognized by Key, Odegaard, and others. The conduct of the Kennedy and Nixon presidential campaigns in 1960 bears implicit witness to the seriousness with which professional politicians took its cogency as a practical guide to the ingredients of American voting behavior. To assess the full stature of what Campbell and his colleagues at the University of Michigan Survey Research Center have achieved, however, it remains to note that this work must now be added to the small and select array of classics of substantive social psychology.

American social psychology has consistently underplayed the substantive description and psychological analysis of important areas of social behavior, setting greater store upon developing refined techniques of data collection and measurement, on the one hand, or discovering general process laws, on the other. At the beginnings of an era of social psychological research, we find Thurstone's ringing title, "Attitudes Can Be Measured" (1928)—launching a tradition in which the logic of measurement took priority over interest in the phenomena under study. The other emphasis, on general processes and relationships, may be exemplified by the theoretically focused lines of development initiated by Lewin in the study of group dynamics and those carried forward at Yale and at Michigan in the study of attitude change.

Social psychologists may justly take pride in these achievements. Both kinds of emphases potentially lay the ground for the systematic description and analysis of social life in its major facets from a social psychological perspective, but the latter task has to be attacked in its own right. To attack it profitably, moreover, takes substantial resources. Our methodological advances and our aspirations for a cumulative science put the damper on attempts in this direction from the psychological armchair—though products of such endeavors by an earlier generation still make stimulating reading, and we find ourselves surreptitiously turning outside the psychological fraternity to social critics like Riesman to fill the present lack. Nor can the ever-available college sophomore tell us what we need to know. Small wonder, then, that there are few monographic classics in this vein to put on the shelf beside *The Authoritarian Personality* and *The American Soldier*. The Survey Research Center has put its special resources to good use in producing one.

The book is based primarily on a two-wave interview survey of a national cross-section at the time of the 1956 election. But it also draws on data from a major study of the 1952 election (previously reported by Campbell, Gurin, and Miller, *The Voter Decides*, 1954), and from a less extensive national survey in 1948, thus effectively funding the experience of the Survey Research Center in the study of presidential voting. Unlike many survey-based reports, it is organized around a theoretical analysis of the voting decision, not around the design of the Michigan surveys that provide its data. It is very well written: there are no tell-tale signs that four authors from two different disciplines had a hand in it. My experience indicates that advanced undergraduates can read the book profitably—if given some help; but the rich and closely analyzed material requires careful and attentive reading.

In contrast to the sociological emphasis of Lazarsfeld and his Columbia colleagues in their community-based election studies (Lazarsfeld, Berelson, and Gaudet, 1948; Berelson, Lazarsfeld, and McPhee, 1954), the Michigan group approaches the national data from an explicitly psychological perspective, in the general spirit of Lewinian field theory. The field theoretical strategy dictated an attempt to link the voter's decision to his immediately preceding political attitudes—in the 1952 study (*The Voter Decides*), party identification, issue orientation, and candidate orientation. Although high correlations were obtained in the earlier studies, this approach was vulnerable to the criticism (cf. P. Rossi in Burdick and Brodbeck, *American Voting Behavior*, 1959, pp. 5-54) that the criterial act and its predictors could be viewed as aspects of the same psychological entity; in Rossi's view, the authors stopped short of addressing themselves to the really interesting questions concerning the sources of these political attitudes.

The American Voter meets this challenge by extending the range of causal analysis. As a framework for the book, the authors develop a conception of what they call the "funnel of causality"—with the act of voting at the mouth or apex, the voter's explicitly political attitudes as immediate determinants in the narrow region contiguous with the mouth, and, as one moves back in time, a host of ramifying prior determinants of these attitudes. These include a core of explicitly political factors and a periphery of nonpolitical social and economic circumstances that require "political translation" before they become causally relevant to the voting decision.

In the present book, party identification is dropped as an immediate attitudinal determinant and placed in the political core that extends in temporal depth back from the mouth of the funnel, because of solid evidence for its early establishment in political socialization and later relative stability. An ingenious method is applied to data for the two Eisenhower elections to assign relative weights to the several attitudes toward candidates, parties, and issues that are taken as immediae determinants; but the major achievements of the book lie in clarifying relationships among factors in other regions of the funnel. The funnel serves, in my view, as a useful scaffolding rather than as a formal model to be accepted or rejected; it is the authors' declaration of strategy in approaching relevant relationships, pragmatically justified by the many instances in which they have contributed notably to our understanding.

What are some of these instances? Space permits notice of only three major contributions among many that could be mentioned, and brief discussion of a fourth topic that is handled less satisfactorily.

Consider first the treatment of political conceptualization and ideology. The detailed analysis of references to domestic issues in free-answer responses to questions about the good and bad points of the two candidates and the two parties exemplifies the "phenomenological approach" to social psychology called for by MacLeod in 1947 but rarely employed so systematically or advantageously. Of the members of the sample who said they voted in the 1956 election, only 15% gave responses that could be classified by the most generous interpretation as based on ideological principles. If the 1956 data are representative, political analysts who interpret election results as reflecting shifts in the "popular mandate" along a left-right, liberal-conservative continuum may thus be misapplying the categories of articulate political leadership. Going beyond phenomenology, correlational analysis

that is presented concerning the structure of political attitudes strongly suggests that what masquerades as economic liberalism—say, on the part of voters from the lower socio-economic ranks—is better understood as pocket-book-interest voting from a much less ideological perspective. The marginal shifts in vote that throw an election one way or the other, moreover, seem to occur mostly among voters who are minimally involved and utterly unideological.

The effects on political attitudes of membership in social groupings are analyzed along lines congruent with principles established in the social psychological laboratory. A sophisticated analysis that could only be done with a large and representative sample shows, in the case of union members, Catholics, Jews, and Negroes, that the higher the identification of the individual with the group, the higher the probability that he will think and behave in ways that distinguish members of his group from nonmembers. The authors go on to show, in cases available for test, that as a person's perception of the proximity between a group and the world of politics becomes clearer, his susceptibility to group influence in political affairs increases. But among a person's group memberships, party identification—at the extreme of close proximity to the political—plays a unique role as "a bridge from group identification to 'proper' political behavior," providing a kind of fly-wheel of stability to the system. All of this is worked out in meaningful detail in relation to cogent data.

What of the role of social class? In an original and important treatment of this traditional focus of research on political behavior, the authors start from the ambiguous standing of "class" as a self-conscience group striving toward recognized goals. Their thesis is "that the 'group' reality of the social class is variable. Under certain circumstances, it is not difficult to conceptualize the social class as a 'group.' Under other circumstances, it is hard to see it as more than a vague demographic aggregate, arbitrarily marked off for purposes of analysis." Where the group model fits, they show, the principles found to operate with other social groupings also hold for class. Via an index of "status polarization" the authors explore short term fluctuations in the tendency toward class voting, and social and psychological factors that tend to promote and to limit it. Party identification once more emerges as a conservative influence on those who are strongly identified, in periods (such as the Eisenhower elections) when status polarization is in decline, or in other periods when it is rising.

Although *The American Voter* devotes a chapter to personality factors in voting behavior, its contribution in this sphere is less satisfactory. The major investment of the 1956 election study in this respect was inclusion of items from Christie's balanced version of the F scale, and with acquiescent response set controlled, no intelligible relationships with issue responses emerged. On theoretical grounds, the authors opt for a narrow or residual conception of personality, more or less restricted to "deep" motivational factors. Such a conception minimizes the likelihood that signficant relationships between personality and political behavior will be found. If a more extended view of personality is entertained, to include a person's major cognitive traits (surely related to level of political conceptualization), his temperamental traits, values, and identifications, the story might well be different.

But this is to carp at a minor aspect of a book that in whole and part is otherwise very satisfactory indeed. It is far and away the best existing analysis of American presidential voting; I would consider it the most substantial contribution yet made

from the behavioral sciences in the political area. And I can think of no piece of survey research that makes a more impressive psychological contribution. Limited though its central data are to the Eisenhower elections with all their peculiarities and time-boundedness, the book is informed by a comparative and historical perspective that establishes its valid claim to more enduring significance. As for the questions that remain open, and there are many important ones, it is heartening to know that the Survey Research Center's program of political studies is continuing. We can anticipate sequel volumes, and our anticipations of them will be high.

University of California, Berkeley
M. BREWSTER SMITH

The Manipulation of Human Behavior. By ALBERT D. BIDERMAN and HERBERT ZIMMER (Editors). New York, Wiley, 1961. 323 pages, $7.95.

Few questions of applied psychology have aroused as much interest and been surrounded by as much fiction and fantasy as the question of how much power an unscrupulous interrogator really has over an unwilling subject. We think of him not only as using the classical torture devices, but also as a skilled hypnotist, psychopharmacologist, neurophysiologist, sensory depriver, and social psychologist. Should we wonder that the layman often shrinks back from a modern psychology which promises so much control over mind and behavior?

The question, of course, is how much of this image is fact and how much is fancy? What areas of behavior are subject to manipulation, even in an unwilling subject, and what are the probabilities of successful resistance? Biderman and Zimmer asked a number of contributors to discuss various psychological and physiological techniques which might be brought to bear in interrogation. The result is a fascinating and tremendously useful review of psychophysiology, sensory deprivation, psychopharmacology, lie detection, and hypnosis. An additional chapter on subcortical stimulation by implanted electrodes was commissioned but not completed because the reviewer, Dr. Sidney Marvin, felt that the known methods were too crude and experimental results too tenuous to warrant its serious consideration as a practical method of manipulation.

This collection will be of interest both to the practitioner (i.e. the police interrogator) and to the experimental psychologist. It brings together under one cover a tremendous amount of information, and presents it at a level of complexity appropriate to the material. The individual authors do not attempt to write popularized extrapolations, but rather, serious, considered reviews and analyses of the issues. At the same time, the practitioner will find neither the technical vocabulary nor the discussion of theoretical issues to be obscure or irrelevant. Most of the chapters are well organized, making it possible for the reader to discriminate easily the more or less technical parts of a chapter.

Looking at the volume as a whole, what is the answer to the question of human manipulability? One can capture the flavor of the individual conclusions best simply by saying "it all depends"; in most instances it depends on those elusive and stubborn variables—the social situation and the personality. Thus, certain drugs can be very effective in eliciting information, but not in all situations or with all people; and there is the uncomfortable possibility (for the psychopharmacologist) that those who give information under drugs are the same ones who would give it without drugs. Is it worth addicting a subject in order to gain control over him when com-

parable control can be established by the withholding of water? Sensory deprivation has a reliably established disorganizing effect on human functioning, but whether such disorganization also leads to willingness to reveal protected information is far from established. Again, it seems to depend somewhat on the personality and situation. Lie detection can be quite accurate in picking up the lying subject, but not in all people—apparently some can chronically fool the machine and the operator. What of hypnosis? It seems well established according to Orne that one cannot induce trance in a resistant subject, unless there is a positive relationship between him and the hypnotist. Whether the subject will then behave in a manner inconsistent with his own standards again seems to depend on his relationship with the hypnotist, and we again learn that the subject who would talk under hypnosis is probably the same one who would talk without being put into trance.

A second theme recurs in a number of the reviews. The more you debilitate, drug, perceptually deprive, etc. the subject, the more confused he becomes and the more unreliable his statements become. Thus the interrogator must always balance the gains derived from manipulative interventions against the risks thereby incurred of obtaining fantasy or confabulated material. Subjects wishing to protect certain items of information seem to be able to do so even under hypnosis or under the influence of drugs.

The third and perhaps most important conclusion reached by each reviewer is that the state of the art appears to be well ahead of the state of the science. Especially in the fields of hypnosis, psychopharmacology, and sensory deprivation there are insufficient well-controlled studies to make possible reliable inferences concerning the effects of such operations on the motives and thought processes of individual subjects. If this book succeeds in establishing clearly only this one fact, it will have been most worthwhile, if only to destroy once and for all the stereotype of the psychologist-interrogator as omniscient and omnipotent.

The individual chapters are of somewhat uneven quality. Gottschalk's review of the uses of drugs and Orne's review of the uses of hypnosis are the most pointed toward the problem of interrogation, the most clearly organized, and conceptually the most stimulating. Hinkle's discussion of the "brain syndrome" is in spots too conjectural and oversimplified. It is hard to believe that the whole range of stress conditions he discusses all lead essentially to one kind of syndrome, and that the psychosomatic point of view can be so safely ignored. Hinkle asserts that the brain syndrome involves increased suggestibility but does not substantiate this with experimental data.

Kubzansky's review of studies of reduced environmental stimulation is excellent and much needed. It falls short only in the one respect that it assumes the interrogation situation to involve more *sensory* deprivation than it actually does. It is only now beginning to be recognized that it is *social* under- or over-stimulation which is the critical variable both in interrogation and indoctrination. Future studies must pay more attention to identifying the particular effects of social isolation and must study also the phenomenon of over-stimulation. Davis' review of lie detection makes painfully clear the limited number of experimental studies from which one can make valid inferences.

The weakest chapter is Blake and Mouton's review of interpersonal influence because it is organized in terms of a theoretical scheme which is unsuited for extrapolation to interrogation. Thus we find a workmanlike discussion of over 100

experiments in social influence, but little attempt to select the more or less important ones in terms of implications for interrogation. And in the process of fitting experiments into their scheme, Blake and Mouton have missed some (e.g. Rosenberg's studies of attitude structure) which have direct implications for an understanding of how and why some interrogators are as successful as they are. It is regrettable that the field of social psychology which has perhaps the most to offer to an understanding of interrogation should appear to be as impoverished as this review implies by its content and conclusions.

The reviewer also wishes that the editors had seen fit to include some of their own material, notably an excellent socio-psychological analysis of the interrogation situation and its face-to-face demands by Biderman (*Sociometry*, 1960, 23, 120-147), to indicate how powerful are the interpersonal forces which an interrogator has available to him, thus making unnecessary a lot of pharmacological or hypnotic manipulation.

Finally, a word must be said about Meltzer's excellent closing chapter, on malingering as a way of countermanipulating by the subject. Meltzer's discussion makes it quite clear where such an approach is likely to succeed and where to fail, and reveals beautifully how much the interrogation situation is a game of wits between two people in which all the odds favor the interrogator because of his greater control over the situation.

University of Texas JACK CAPEHART

Experimental Design in Psychological Research (Rev. ed.). By ALLEN L. EDWARDS. New York, Rinehart and Co., 1960. Pp. xiii, 398. $6.50.

The original edition (1950) of *Experimental Design in Psychological Research* is probably as well known to psychologists as any book in the field. Those who are familiar with the first edition will quickly recognize the family resemblances in the revision. Though there have been some changes in content and emphasis in ten years, writing style, pedagogical technique, and the model for intra-chapter organization are basically the same.

About one-fifth of the revised edition is new material. The major additions cover analysis of variance models, multiple comparisons, and techniques for applying orthogonal weighting coefficients to designs ranging from the simple randomized to trend analysis. The new book is shorter than its predecessor. Discussions of *chi square, t, r,* scale transformations, and heterogeneity of variance have been abbreviated. Nevertheless, at least one-third of the book is still devoted to topics that would normally be covered in a first undergraduate course. The present review will be concerned principally with the new material in the book and with the general approach to teaching experimental design.

The treatment of analysis of variance models is somewhat disappointing. Rather than incorporating each model with the discussion of the associated test, all models are considered in a single postscript chapter. The effect is to underplay the importance of models in the design and analysis of experiments. Furthermore, the term *model* is not defined, and none of the models discussed is ever explicitly stated. Tables of expected mean squares are included for a number of designs, but the value and meaning of these tables is diminished by the failure to define *expected mean square.* At one point the ratio of two expected mean squares is incorrectly shown as equal to *F.* Consequently a student might conclude that expected

mean squares are statistics with sampling distributions rather than parameters describing sampling distributions. Rather than rationalizing or analytically determining expected mean squares for the models discussed, the treatment of this topic consists largely of mechanical rules for generating expected mean squares. Some of these rules are reduced to modifications of the subscripts of the previous model.

Experimental examples are presented to illustrate two of the mixed models. The choice of variables in one of these examples is unfortunate, though correct. In the study two experimenters are used, one male and one female. *Sex* is treated as a fixed variable in the model. It is not made perfectly clear, however, that generalizations are restricted to the particular experimenters used, and that samples of size one leave the sex variable hopelessly confounded with other differences between the two experimenters.

The new material on procedures for making specific and analytical comparisons among means is a welcome addition. Duncan's "New Multiple Range Test," Dunnett's test for comparisons with a control group, and Scheffé's test are all included in a chapter on multiple comparisons. Examples are given to illustrate the computations required for each test. The discussion of the general characteristics of such tests is rather limited. More space might well have been allotted to questions of when the tests should be used and to the notion of error rates.

Procedures for making orthogonal comparisons are also introduced in the chapter on multiple comparisons. The treatment of this topic is also mechanical, the emphasis being on computational techniques. Except for some errors of notation, examples are straightforward and easy to follow. However, there is reason to wonder whether the student who computes a sum of squares for a comparison will understand what he has done. Magical and mystical formulas are left magical and mystical. In view of the importance of orthogonal weighting coefficients, it would have been worthwhile to show that the sum of squares for a comparison is simply the product of a squared correlation coefficient and the total sum of squares. The correlation measures the relationship between a function of the independent variable, described by the coefficients, and the values of the individuals on the dependent variable.

The chapter on trend analysis deals with studies in which one or more groups of *S*s are measured over trials. (In an earlier section the term *trend analysis* is used to describe simple single classification designs involving an ordered independent variable.) In the initial half of the chapter examples and analyses are set forth much as they were in the earlier edition. In the second half new techniques are applied to the same examples to show how the sums of squares for over-all, group, and individual trends can be partitioned into linear, quadratic, and higher order components. Though the sequence of development of this material is reasonable, one might contend that the chapter should have begun where it ended, with a consideration of individual trends. The content of the chapter falls neatly into place once it is understood how the learning curve for a single *S* is described for such analyses. The descriptive variables, all independent, are functions of (1) the average score over all trials, (2) the slope of the curve, (3) the quadratic component to the curve, etc. From this point it is easy to conceptualize a trend analysis as a series of analyses comparing groups on each of the descriptive variables.

In some respects a better title for Edwards' book might be *Experimental Designs in Psychological Research.* Like the earlier edition the book deals with a number of specific experimental plans, computations for them and expositions of some

considerations involved in their application. Though principles of experimental design are occasionally discussed, they are not heavily emphasized.

It is easy to argue that an author should have written a different book, that he should have included some topics and omitted others. While the prejudice of the reviewers is for a book which would eliminate much of the elementary material of the first 100 pages, which would give more space to a number of advanced topics, and which would emphasize principles more heavily, many instructors will undoubtedly find the revised edition of *Experimental Design in Psychological Research* a very useful text. Despite its shortcomings, the book it more readable and more comprehensible than most of its competitors.

University of Arizona ROBERT E. MORIN

Psychological Scaling: Theory and Applications. Edited by HAROLD GULLIKSEN and SAMUEL MESSICK. New York, John Wiley and Sons, 1960. Pp. xvi, 211.

That the application of mathematical models to treatment of psychological data is an integral part of both substantative and methodological aspects of much current psychological thinking can no longer be doubted when one studies the fourteen chapters appearing in *Psychological Scaling.* Based on the written contributions of fourteen of twenty-two distinguished participants at a conference upon psychological scaling held in Princeton, New Jersey in May, 1958, the contents of the volume will be of interest primarily to members of Division 3 and Division 5 of the American Psychological Association, although a large proportion of them will find it necessary to do a great deal of preparatory reading (e.g., Guilford,[1] and Gulliksen,[2] and Torgerson[3]) in order to grasp the impact of most of the chapters. How far psychometric theory has progressed since the appearance in 1936 of Guilford's *Psychometric Methods!*

Organized about five topics, each of which constituted the context for discussion during each of five different sessions, the chapters, subsequent to an introductory one by Gulliksen, also form five logical groupings: (1) *scaling properties,* comprising Chapters 2 through 5, in which characteristics of category scales and quantitative estimation scales are discussed by Lyle Jones, Torgerson, Shepard, and Green with particular reference to their implications for the nature of psychological judgments; (2) *psychophysical scaling,* treated in Chapters 6 and 7, in which properties of various types of scales (such as ratio, category, and confusion) and the loudness function are considered by Stevens and by McGill respectively; (3) *test theory and psychological scaling,* found in Chapters 8 and 9, in which Lazarsfeld outlines his thinking upon the relationship of latent structure analysis to test theory and Lord, in equating ability to a latent continuum, proposes means for inferring a true score; (4) *choice and the measurement of utility,* covered in Chapters 10 and 11, in which Ward Edwards describes additive and non-additive models in the maximization of subjectively expected utilities and their implications for measurement of value, and Siegel sets forth an axiomatic basis for higher ordered mettric scaling

[1] J. P. Guilford, *Psychometric Methods.* (2nd ed.). 1954.
[2] H. Gulliksen, Mathematical solutions to psychological problems. *Amer. Scientist,* 1959, 47, 178-201.
[3] W. S. Torgerson, *Theory and Methods of Scaling,* 1958.

and evaluates his model in individual decision-making; and (5) *multivariate scaling models,* constituting the basis for Chapters 12 through 14, in which Coombs, Tucker, and Abelson individually treat varied aspects of multidimensional scaling.

Throughout this broad survey of recent thinking in scaling theory numerous applications are cited with respect to measurement of sensory magnitudes, cognitive abilities, utility, and attitudes. Several of the examples, however, constitute rather specialized illustrations, the immediate utility of which many an applied psychologist and even research worker in industrial and educational psychology would find difficult to see. All in all, the volume affords both a broad perspective of recent developments in scaling theory and an appropriate balance in the types of models and sorts of applications being made.

Of encouragement to the reader with somewhat limited mathematical training is the fact that most of the major conceptual material is expressed in verbal terms with a minimum of mathematical notation. In his chapter concerning intra-individual and inter-individual multidimensionality Tucker has wdely deferred to the end an extensive mathematical outline of his presentation.

Despite the diversification of interests and emphases of the writers the volume is well organized and well written. Particularly helpful and pleasing are the numerous captions that serve to guide the reader in his grasp of countless concepts that could readily be lost in a less suitable format. Although by no means exhaustive, the bibliography of more than 150 references does include references covering much if not most of the significant work done during the past fifteen years in scaling and in theory of measurement.

In short, the reviewer found the volume to be informative, well organized and clearly written, representative of current research thinking in psychological scaling, and useful as a reference source. Although highly specialized in coverage, it should find its way to the shelves of most students and research psychologists who are concerned with scaling theory, sensory psychology, decision making and choice behavior, and factor analysis. Its excellence should stimulate many to think about several of the problems posed, to formulate additional problems, and to make important original contributions to their solution.

University of Southern California

WILLIAM B. MICHAEL

Meaningfulness and Verbal Learning. By BENTON J. UNDERWOOD and RUDOLPH W. SCHULZ. Chicago, J. B. Lippincott, 1960. Pp. vi, 430. $6.50.

This book is primarily a report of a three-year research program concerned with the analysis of the fundamental attributes underlying the well-documented effects of "meaningfulness" on verbal rote learning. It consists of 16 major experiments, in which a reported 1436 subjects were presented with 648 items for learning 681,720 times, plus additional ratings of over 1600 items by 1200 subjects for various characteristics (*e.g.* frequency, pronounciability) related to meaningfulness. Much of the latter material is included in a 120-page appendix, which represents the most complete source of scaled verbal material now available, and should extend the range of interests in this book to all researchers who use verbal material for any purpose. The net result may well become the most important book in its field since Ebbinghaus.

The presentation enables the reader to follow, step by step, the authors' pursuit of the nature of meaningfulness, and thereby provides a detailed case-history of the hard-headed systematic empirical approach long advocated and exemplified by Under-

wood. The numerous pitfalls and setbacks encountered along the way receive their full share of attention, and the amount of effort directed toward side issues only indirectly relevant to the main problem is both extraordinary and commendable. Unfortunately, this attention to detail also complicates the task of the reader, especially in the last half of the book where both organization and quality of writing deteriorate somewhat from their initially high standard. To maintain maximal continuity, the reviewer would suggest either reading Chapter 9 on Letter-Sequence-Habits between Chapters 5 and 6, or omitting it entirely on a first reading.

In view of its sad fate at the hands of several experiments, the tenacity with which the authors still cling to their "spew" hypothesis (Meaningfulness = Frequency of experience with material) at book's end may come as a surprise. Due to the high predictive accuracy of the "pronunciability" attribute, the burden of explanation necessarily gets shifted from printed to emitted frequency, on which data are admittedly sparse. Consequently, the reader is likely to feel let down by this concluding interpretation, considering the voluminous research effort preceding it. This, however, is in large measure the fault of the problem, which requires the disentanglement of several highly-correlated attributes without adequate tools for the job, since none of these attributes can be scaled with sufficiently high reliability. Viewed in this light, the extent of the authors' advance toward a resolution of the problem becomes far more impressive.

Although the range of content is limited, the authors have much to say to researchers both within and outside the confines of verbal learning. The usefulness of a 'stage' analysis of verbal learning is conclusively demonstrated, as is the value of combining correlational with experimental methods of analysis. In the final chapter, they join the growing chorus pleading for more attention to the verbal (introspective?) reports of subjects in learning experiments. Nevertheless, the book's uniqueness as a many-faceted systematic attack on an exceedingly complex problem remains its outstanding feature. If it is at all representative of future trends in verbal-learning research, the currently popular characterizations of that area as sterile, artificial, or rigid are about to become distinctly outmoded.

University of Virginia — William F. Battig

Automation and the Worker: A Study of Social Change in Power Plants. By Floyd C. Mann and L. Richard Hoffman. New York, Henry Holt and Company, 1960. Pp. xiv, 272. $4.50.

This is at once a superior piece of research in industrial psychology and a reminder to psychologists that it is extremely difficult to attain tight research designs in an ongoing industrial operation.

The authors have provided us with some valuable data on the behavior and attitudes of workers in modern electric power plants. They have found some differences between workers in a highly advanced plant, with automatic feed-back controls and other technical innovations, and workers in a traditional plant. Unfortunately, it is virtually impossible to tell what proportion of these differences may be ascribed to the automated equipment, and how much to other variables.

Workers are better satisfied in the new plant; but their wages are also higher than in the old one, and wages notoriously have some effect on job satisfaction. Foremen have more authority, do not perceive their role as "messenger boys for top management" in the new plant; but the top plant executives in the new plant endorse a philosophy of decentralized control, whereas top management in the old plant held

to tight central controls. Does automation change top managerial philosophy? It seems unlikely. Hence the behavior of the foremen (of which the workers approve) may be a function of organizational climate rather than of automation. The workers enjoy their job roles, which are more varied and interesting, but they also have a chance to socialize more on the job. (In C. R. Walker's *Toward the Automatic Factory,* workers complained of automated jobs because they were relatively isolated from their fellows.) Finally, the workers in the new plant were not a random sample of the company's prior employees; by some curious chance, they show a much higher percentage of Republicans, and they voted against a union to represent them, whereas all the company's other plants were unionized.

It is thus unjustified for the authors to say, "Our present findings emphasize the profound effects which the design and location of equipment have both on the various jobs which workers perform in a plant and on the kind of organizational and group relationships which develop among them" (p. 193). The findings show differences which, as noted above, are open to various interpretations. The problems are not due to inadequate research design, but to the fact that in industry precise controls simply cannot be observed in most investigations.

In a concluding chapter on "administrative and research implications" the authors do an exceptionally fine job of identifying problems which management should face before embarking on automation, and they speculate intelligently on the role of the technological factors *per se* in producing their data. Many of these insights derive from casual interviews with workers and from on-the-spot observation, rather than from the questionaries of the Survey Research Center. The latter, while admirably constructed, could not pick up some kinds of crucial data. This is not to deprecate the use of questionaries but to suggest that they are most useful when neatly embedded in a careful study of the history and overall functioning of the organizations within which they are employed.

This book is highly recommended for industrial psychologists and for industrial executives.

Wayne State University

ROSS STAGNER

The Training of Psychotherapists. Edited by NICHOLAS P. DELLIS and HERBERT K. STONE. Baton Rouge, La., Louisiana State University Press, 1960. Pp. xii, 188. $5.00.

This is a thoughtful discussion by 16 authorities (seven psychiatrists, seven psychologists and two social workers) on psychotherapy, what it is, who should do it, and what training is to be desired. The book, a report of a symposium held at LSU, begins with a provocative discussion by August Hollingshead of his research on social class and mental illness. He concludes that "latent social factors, besides claimed medical criteria" help to determine the extent and nature of psychiatric treatment. Other discussants consider what kinds of training best equip an individual to perform psychotherapy. Academic psychologists will be impressed that much of the discussion is about topics that graduate schools are not inclined to consider worthy of academic credit, but are currently considered a part of internship or practicum training.

The reviewer was intrigued that two of the discussants reminded the members that there was no empirical evidence to indicate that a psychotherapist who had undergone personal psychotherapy was a better psychotherapist than one who had

not had personal therapy. Despite this, all the other panelists agreed that personal therapy was a decided advantage, and some insisted that it was a basic essential. Although some of the panelists are psychoanalysts, there appeared to be almost unanimous agreement that psychoanalysis should no longer be considered the treatment of choice with a majority of emotional problems.

Psychologist Gordon Derner won approval of his plea that future psychotherapists should be trained broadly and in such areas as behavioral science, social science, anthropology, "the biological and physical aspects of man," philosophy and mythology. The group rejected the proposal made elsewhere by Kubie that there be a six year training program to combine the useful aspects of medical training with training in psychology. The general theme seemed to be broad training and flexibility of approach to the patient, according to the particular patient's presenting problems.

Academicians will find the discussion interesting but difficult to translate into semester hours of credit for this or that topic. As a matter of fact, admitted by the editors, the topic of training received less attention than did the question of what is psychotherapy and who should do psychotherapy. Perhaps this indicates that problems of professional role are still uppermost in the minds of clinical psychologists and psychiatrists who are on the "firing line" in clinics and hospitals. At least one participant has the role of psychiatric resident and psychological intern neatly differentiated: "The residents tended to meet crises more with authority and directive action, while psychology trainees tended to meet them more with withdrawal and detachment."

University of Richmond AUSTIN E. GRIGG

Transactions: Fifth Conference on Neuropharmacology, Princeton, New Jersey, 1959. Edited by HAROLD A. ABRAMSON. New York, Josiah Macy, Jr. Foundation, 1960. Pp. 251. $6.00.

Four presentations make up this fifth and final conference in the neuropharmacology series which was sponsored by The Josiah Macy, Jr., Foundation. Udenfriend takes the interesting tack of comparing the metabolism of the entire group of neurologically-active amines, among which are epinephrine, norepinephrine, acetylocholine, serotonin and histamine. Ordinarily considered separately, these substances turn out in this very useful discussion to share common features of origin and mode of metabolism. This presentation and accompanying discussion are biochemical in orientation and will be valuable to anyone concerned with the intimate details of what happens to these substances in the body.

The central actions of chlorpromazine and reserpine are discussed by K. F. and Eva Killam, and by Brodie. It is by now clear that the common tranquilizing effects of these drugs are produced by markedly different physiological mechanisms. The Killams point this up in two ways. First, they show that neither drug does the obvious and suppresses or inhibits the reticular formation. How reserpine acts is not clear, but chlorpromazine seems to produce its effect by enhancing the inhibitory effects of the reticular formation on incoming sensory impulses. Secondly, they show that these drugs have different effects on electrical brain rhythms of cats fully trained in a conditioned avoidance response. In these investigations, carried out with E. R. John, reserpine brought about what they call a "detraining" or "decoupling" of the rhythms to a level observed during acquisition. Chlorpromazine on the other

hand seemed to produce its effects by interfering with transmission of afferent impulses from the conditioned stimulus.

Brodie develops his hypothesis of how the two drugs act. The focus of action for both drugs is the hypothalamus. Chlorpromazine is hypothesized to block a central adrenergic system related to Hess's ergotropic system while reserpine acts by stimulating a serotonin system related to the trophotropic system. Several new analogues of psychotropic drugs worth noting are mentioned in the presentation. Koella presents data to show that three separate factors are concerned in how serotonin depresses cortical potentials evoked by light flashed into the retina. Neural impulses from the carotid sinus, a bilateral inhibition from subcortical levels, and an ipsilateral cortical inhibition are all influential.

These transactions are essentially verbatim accounts of the spontaneous give-and-take at the conference. Their value has been materially enhanced by including expert summaries of the discussions by Quastel and by Pfeiffer. The reader would be wise to sample the expert offerings in this volume by beginning with the summaries.

Cornell University ALLAN C. GOLDSTEIN

Learning and Behavior. By REED LAWSON. New York, The Macmillan Company, 1960. Pp. xii, 447. $6.75.

This is an undergraduate-level text aimed at the usual course in learning taught from a behavioristic, empirical viewpoint. The book is organized around a "stimulus-response-reinforcement" model of learning phenomena into which most such phenomena have been forced, often in a quite original and constructive way. The technique of the author is to describe experimental paradigms in terms of variations in one or more of the three aspects of the model. This organization in combination with the sometimes idiosyncratic terminology used in the book results in chapter subtitles which may irk the purist such as "One R/ Motivation Relevant/ X Not Pertinent, 'Acquired.' " From the point of view of the student, the strong thread of organization tends to tie together a mass of material which in other books is often seen as a diverse conglomeration of facts and theories.

An interesting feature of the book is the relatively liberal use of extended quotations from original sources when they make Lawson's point. These quotations and the references are carefully selected as illustrative studies. The book is not a catalog of names and experiments. As a result, Lawson's list of references is only about half as long as that of another very recent book dealing with the same material at a comparable level.

In the first of twelve chapters, the author outlines his model, discusses a number of terms, and justifies the behavioristic, comparative approach. The next four chapters deal with "simple habit formation," discussing Pavlovian and operant conditioning, characteristics of reinforcers, and the effects of reinforcement variations in operant conditioning. Then follow chapters on discrimination learning, transfer of training, and "eliminating (changing) simple habits." The next two chapters deal with learning situations that do not seem to fit the model (*e.g.* place learning, sensory preconditioning) and with relating motivational concepts to his admittedly incomplete model. Finally the book deals cursorily with the acquisition of habits in complex situations (*e.g.* rote learning).

The book seems particularly well-suited as an introduction to the theories and facts of the field of learning which derive from the work of Hull and Skinner and

their adversaries. The strong point of the book is its implicit assumption that the undergraduate is taught best when exposed only to selected, representative experiments which highlight the various phenomena and theories of the field in a well-organized context. The disadvantage of the approach is that it makes things seem simpler than they are.

Duke University C. ALAN BONEAU

Management in Marketing. By HECTOR LAZO and ARNOLD CORBIN. New York, McGraw-Hill Book Company, Inc., 1961. Pp. xviii, 657. $8.50.

In recent years increasing attention has been paid to what has been referred to as "the new marketing concept." Usually this means a movement away from the simple concept of marketing, which was conceived of as selling anything the factory could produce, to the more current concept, which is customer-oriented for profit. Under this new approach, all activities of a company which are not strictly financial, manufacturing, or technical, are integrated under a single marketing head and oriented toward customers and customer-needs. Drucker in his short forward expresses this well when he says, "Economically, all activities of a business other than marketing are 'costs' (or at best 'risks'); only marketing produces 'results.'"

The material presented in this book is based on three years of research, two national mail surveys covering more than 1000 business firms, several dozen personal interviews and conferences with business leaders, writers, and management consulting executives.

The content of the book is presented in three parts: Part I, Preliminary Phases, includes chapters mostly concerned with organizing for marketing; Part II, Management in Specific Marketing Functions, deals chiefly with chapters on marketing research, product planning, sales and advertising, and distributions; Part III, Elements of Managerial Administration, includes chapters on control in marketing, people in marketing, and benefits of marketing management. Each part is followed with case studies illustrating the material presented. The book also contains an appendix and index.

According to the authors, many economists and business observers believe that the decade of the sixties is destined to become the decade of marketing in the same sense that the decade of 1900 to 1910 was the decade of finance and the decade of 1910 to 1920 was the decade of production. There are some signs indicating this kind of a trend.

It is the opinion of the reviewer, however, that there is more to business than marketing and there is more to organized thinking about business problems than marketing concepts, particularly when human organization in industry is as neglected as it is in this volume. But, as good books on marketing go, this is a relatively comprehensive book, practical, useful for people in the field, and it can do some good for some people interested in the field because of their relationship to the implications of the work.

Human Relations Research Foundation H. MELTZER
St. Louis, Missouri

Stuttering and What You Can Do About It. By WENDELL JOHNSON. Minneapolis, University of Minnesota Press, 1961. Pp. ix, 208. $3.95.

Having completed some thirty years of work on the problem of stuttering, a noted authority on speech disorders writes a book addressed directly to stutterers and

parents of stutterers. Practical suggestions for help are skillfully interwoven with the story of the author's efforts to cope with his own stuttering, his years of research on the problem at the University of Iowa, and the development of his present views on the subject.

Johnson's vision of the truth about stuttering is a seductive one. He has, moreover, a gift for communicating his excitement about it. Consequently, the layman will have to read this book very carefully to discover that there is much doubt and controversy among professional workers regarding the etiology and treatment of stuttering. The basic message contained in this book is that at the moment the problem of stuttering first arises for a child's parents his nonfluencies are essentially similar to those of most other young children. They begin to differ from normal nonfluency only when he has reacted for some time to his parents' anxious, perfectionistic misevaluation of his hesitancies. Accordingly, the best way for parents to prevent the problem from developing further, or to treat it in its early stages, is to re-orient themselves perceptively and evaluatively to the child's speech. This thesis is conveyed to the reader with appropriate tact, and supported by parents' descriptions of the earliest speech symptoms which they had identified as stuttering in their children. A notable feature of the book is the inclusion of transcripts of some of Johnson's interviews with stutterers' parents. This material, presented for the first time, provides fascinating insights into the process by which the author draws the particular inferences about stuttering that he does.

As a popularization of a technical subject, this book is far superior to the usual product of the professional popular science writers, and parts of it come close to making a literary genre of this kind of writing. Altogether, it is an absorbing, delightful and, to a specialist who does not accept all of Johnson's basic premises, at times an exasperating little book.

Brooklyn College

OLIVER BLOODSTEIN

Psychology of Literature: A Study of Alienation and Tragedy. By RALPH J. HALLMAN. New York, Philosophical Library, 1961. Pp. 262. $4.75.

There are as many schools of literary analysis as there are of psychology, but Hallman, in this attempt at a theory of tragedy, puts his trust in psychological analysis. His reason is that "tragedy appears to be something other than an art form; unlike most of the other arts, it tries to overrreach itself and to speak about the human situation as it occurs outside of literature." Thus this book is a psychology, not of the reader, nor of the writer, but of a persistent theme, the aesthetic expression of the tragedy of the human situation.

Hallman rejects the classical theory of the tragic flaw in the hero's character and the claim that the tragic experience can occur only in a universe controlled by moral forces. Tragedy began with the Dionysian rituals celebrating the death and return of life-giving vegetation but, unlike primitive man, who identified in ritual with the cosmos, the tragic hero of literature is alienated from his tradition. He knows that death is meaningless. Reason transforms man into a rebel, and what he rebels against are rational constructions that threaten his freedom: a sociopolitical order created by man and a metaphysical order created by God. The tragic flaw, according to the author, is not in the hero but in the human condition. It is not that the universe is moral but that it is irrational. Only by abandoning reason

and regressing to the primitive condition, in the womb of Mother Nature, that is by dying, can he be reborn.

This book must be evaluated by the standards of literary criticism rather than those of empirical psychology, with which it makes only occasional contact. Certainly reading it is a challenging intellectual engagement. "Tragedy forces man to consider the boundaries of his own existence." Knowledge is fragmentizing, we are told, while the "wisdom of the primitive mind, as expressed in its mythology, teaches the essential unity of experience." "Individual thinking . . . arises out of a floundering and failing ritual." Some psychologists will be carried away by these mystical insights into the death drama; few will be reborn.

Michigan State University — DONALD M. JOHNSON

The Problems of Perception. By R. J. HIRST. London, Allen and Unwin; New York, Macmillan, 1959. Pp. 330. $5.25.

British philosophers have for a long time been giving us admirably written, closely reasoned analyses of the processes of cognition. One thinks, for instance, of Bertrand Russell's *Analysis of Mind* (1931), H. H. Price's *Perception* (1933), and Brand Blanshard's magnificent *The Nature of Thought* (1939). The philosophers may seem to dispose of our cherished psychological facts with disturbing ease, but, after all, they are solving problems which we psychologists are only dimly aware of; and, at any rate, they are always a delight to read.

Professor Hirst, a lecturer in logic in the University of Glasgow, writes in the best British tradition. The psychologist might say that his analysis is 'logical' rather than 'psychological,' but he differs from most of his predecessors in that he seems to have some acquaintance with the experimental literature that has accumulated during the past 100 years. Not a deep acquaintance, it is true, for he has virtually ignored the wealth of information that has come from the German laboratories (how can one write on perception without having read Metzger's *Gesetze des Sehens?*) ; but at least he knows that there have been researches into the problems of phenomenal constancy, which Price did not seem to know in 1933, and he is alive to the recent challenges from neurophysiology.

This is not the place for a critique of Hirst's final position, to which he comes after having examined and refuted most of the standard theories. It is a form of Identity Theory which most psychologists will probably find congenial, and certainly not shocking. The reviewer's question is: should the psychologist invest his time in the reading of a book which borders on his own interests but which is essentially philosophical? His judgment is that the investment is a good one. When a philosopher settles down to a bit of clean reasoning about psychological problems, the result is sometimes a thing of beauty. Hirst is not a modern Brentano; but he is very good.

Cornell University — R. B. MACLEOD

Creativity. By EMANUEL F. HAMMER. New York, Random House, 1961. Pp. x, 150. $1.25.

The title of this paper-back is misleadingly broad. Its contents actually deal with the preliminary stages of a study of some very small samples of high-school art students.

Teachers of 18 students holding scholarships selected five whom they regarded as

"truly creative" and five whom they regarded as "merely facile," where 'creativity' meant "communication of real feeling." The author applied four projective procedures, two of which involve the subject's drawing of pictures. He later compared protocols of the two groups.

A kind of cross-validation was made in a second class of similar students, in which 13 were judged in similarly extreme groups, with 11 hits and 2 misses. Additional cross-validations are planned in future classes.

Conclusions are drawn concerning differences in certain traits of "truly creative" vs. "merely facile" art students. As in the case of most results with projective techniques, the descriptions of the trait are often not communicable, nor has validity of the constructs been established. Evidently under some pressure to do so, the author goes well beyond his results in describing creative artistic personality.

Application of the traits in question to creative groups other than artists should be especially tentative. It is doubtful, for example, that scientists are creative because they "communicate their real feelings." It can even be questioned whether communication of feeling, whether "real" or not, is even the best conception of creativity for artists. There is probably a genuine distinction between being expressive and being inventive. Expression of feeling would lend itself better than inventiveness to detection in projective performances, particularly when those performances are in the form of drawings.

University of Southern California

J. P. Guilford

Speech Disorders and Nondirective Therapy: Client-Centered Counseling and Play Therapy. By Robert F. Hejna. New York, The Ronald Press Company, 1960. Pp. vii, 334, $6.50.

This book purports to be a discussion of "client-centered counseling and play therapy" (p. 11) for speech disorders in general. In Chapter 2, entitled "Areas of application," the topics considered are delayed speech, stuttering, voice problems, articulatory disorders, and "other problems." The book consists of 329 pages, all but 100 pages of which consist of transcriptions of the *words spoken* in nondirective counseling and therapy sessions. Practically all of these words are concerned, however, with stuttering, in spite of the promises made to cover other areas.

Apparently Hejna has failed to consider the great heterogeneity of speech disorders. It is as though the psychoanalyst wrote on the treatment of headaches, including under one category all etiologies—worry, eye-strain, concussion, tumor, allergies, etc., and then gave as illustrations of headache therapy only the therapy for "worry" cephalalgia. The present book would be more honestly labeled if called, *Nondirective Therapy in Certain Speech Disorders* or *Stuttering and Nondirective Therapy.*

The book is a good explanation of non-directive therapy (with lengthly illustrations) for those speech cases whom Rogerians consider fit subjects for such therapy.

Brooklyn College

Robert West

INDEX

CHARLES VALLEY BROOK, University of Texas

AUTHORS

(The names af authors of original articles are printed in CAPITALS; of authors of books reviewed, in roman; and of reviewers, in *italics*.)

SUBJECTS

(References in *italic* figures are to reviews.)

Serial vs. paired-associate, 517-528.
Single-trial, evidence for, revaluated, 590-595; tested by free recall, 596-602.